Elucidation of mammary gland development: combination of novel techniques to study proliferation and stem cells

UNIVERSITY OF CAMBRIDGE
Department of Pathology

Olivia Beatrice Harris

Newnham College

September 2018

This dissertation is submitted for the degree of Doctor of Philosophy

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed 60,000 word limit for the School of the Biological Sciences.

Olivia Beatrice Harris

September 2018

The work carried out in collaboration with others is outlined below.

Chapter 3

The optimisations and development of optical clearing methods for the mammary gland was carried out in collaboration with Drs Felicity Davis and Dr Bethan Lloyd Lewis. I was involved in, and contributed substantially to, all aspects of this work. Tumours were provided by Dr Jessica Hitchcock. The research involving the development of human tumour scaffolds and their imaging was carried out by Robert Hume, with the optical clearing being a collaboration between him and myself. Light sheet microscopy of samples was carried out by Dr Matthias Pasche at the MRC Laboratory of Molecular Biology, Cambridge. Optical clearing of the liver was done as a result of a collaboration between myself and Mikel McKie from the laboratory of Dr Meritxell Huch at the Gurdon Institute, University of Cambridge.

Chapter 4

The work involving lineage tracing during embryogenesis, puberty and lactation was all carried out as a collaboration between Dr Felicity Davis, Dr Bethan Lloyd Lewis and myself. This included injections, mouse management, dissections, mammary gland preparation, staining, and optical clearing and imaging. I was involved in, and contributed substantially to, all these aspects of this work. The resulting publications of these date are cited where relevant and attached in the appendix. The tissues from the one year confetti mice lineage tracing into adulthood were a kind gift from the laboratory of Dr Bon Kyong Koo.

Summary

The mammary gland is a dynamic organ that undergoes many cycles of proliferation and death throughout both oestrus cycling and the gestation/lactation/involution cycle. Moreover, it is a unique tissue in that the majority of its development occurs postnatally coincident with the production of ovarian hormones in puberty. This capacity of the mammary gland for rapid growth and regeneration has been attributed to mammary stem cells (MaSCs). However, despite extensive efforts over the past 60 years, definitive characterisation of the stem and progenitor cells of the mammary gland has yet to be achieved. A number of recent conflicting, lineage tracing studies have served only to fuel the fires of controversy, with previous characterisation of the mammary gland largely carried out using two-dimensional tissue sections. However, in order to fully appreciate the capacity of MaSCs and to maintain spatial information of the complex topological structure, the mammary gland must be investigated in its intact form. Accordingly, there is still much disagreement regarding the potency, capacity and location of MaSCs.

Consequently, in order to unequivocally elucidate the MaSC hierarchy, a variety of novel techniques have been combined in this thesis. The first involves development and optimisation of optical tissue clearing techniques to allow the visualisation of the native mammary gland, *in situ*, and in three dimensions. To do so, a number of different optical tissue clearing methods have been assessed and combined with a variety of microscopy techniques to allow the multiple focal planes of the ductal network to be examined. These imaging techniques were then combined with two neutral lineage tracing models; the first, the Rosa26$^{[CA]30}$ model, utilises stochastic continuous clonal labelling to allow for the fate tracking of the progeny from single functional stem and progenitor cells. The second unbiased lineage tracing approach, the Rosa26-Confetti model, allows for the mammary stem and progenitor progeny to be traced with precise timing, with the additional benefit of a multicolour reporter. Next, proliferation was examined in wholemount tissues to investigate the functional requirements for MaSCs, and their potential locations. Finally, these techniques have been combined with an *ex vivo* 3D organoid culture system to investigate the use of culture methods in examining mammary epithelial cell dynamics.

By combination of these techniques, clonally marked regions can be investigated throughout the development of the mammary gland, from the formation of the embryonic mammary

rudiment, to expansion of the ductal tree in puberty, and ultimately their fate in lactation and involution, where the mammary gland fulfils its evolutionary purpose. The mammary gland provides a unique opportunity to investigate epithelial development extra-embryonically that is not available in other tissues. Moreover, study of maintenance and turnover of this organ has important implications for other epithelial systems. Finally, elucidation of the normal MaSC hierarchy also has important implications for understanding the complex heterogeneity of breast cancer and the cell(s) of origin of breast cancer. Given the proposed longevity and suggested ability of MaSCs to survive multiple waves of cell death in involution, they represent a logical candidate for a potential cell of origin of breast cancer.

The work presented in this thesis provides novel insights into MaSCs and progenitors and their potential to contribute to mammary gland development.

Acknowledgments

I would first like to acknowledge my generous funding from the Wellcome Trust, which has allowed all this work to be carried out. They are a fantastic organisation and one I am proud to have been part of.

Science is nothing without collaboration, and this thesis is a testament to this. I would first like to thank my supervisor, Professor Christine Watson. Not only for her insight and wealth of mammary gland knowledge but showing me that it is ok for women in science to wear fabulous shoes. Dr Felicity Davis and Dr Bethan Lloyd Lewis provided a fantastic foundation for not only mammary gland research, but how to be a good scientist in general. It was amazing working and supporting one another through hours in the mouse house and in dark microscopy rooms. They also tried their hardest to keep me safe despite my lack of safety awareness. This is also true for Dr Jessica Hitchcock. She has provided a sounding ear for not only discussions about science, in particular immunology and which is our favourite immune cell is (macrophage of course), but about everything and anything else. Seeing her face every day has kept me going. Ala Alanzi is another lab mate whom I would like to thank. She has shown steadfast resilience whilst also wearing fabulous shoes, even if they are fur boots in the winter rain. I would also like to thank the BSU staff, especially Lisa Wright and Daniel Gates and Peter. They have facilitated all the mouse work which would not be possible without their help. I would also like to acknowledge the many mice used in these studies. Also, Dr Leila Muresan, Dr Kevin O'Holleran and Dr Martin Lenz at the Cambridge Advanced Imaging Centre for microscopy and analysis support.

Finally, the support from my friends and family has been vital for me completing this PhD. These include, John Cassidy, whom has not only provided support but also scientific guidance, through hours of discussion about breasts. Hannah Thompson, whom has shown me what a strong independent woman is. Mikel Mckie, whom is always there to give me reagents when I run out, and Ayla Newton whom has always provided an ear, or distraction with gruesome medical stories. Most importantly, I would like to thank my family. My sisters Molly and Emily Harris, and most importantly my parents Lisa and Jeremy Harris, whom this thesis is dedicated to. They have continuously provided grounding stability and belief that I can do anything. They are too humble to take any credit but I would not be at this point without their support.

Contents

Declaration ..... **i**

Summary ..... **iii**

Acknowledgments ..... **v**

Contents ..... **vi**

Table of figures ..... **xi**

Abbreviations ..... **xv**

Chapter 1

Introduction **1**

1.1. Introduction to the mammary gland *2*

Embryonic mammary development of the mouse 2

Pubertal development of the mouse mammary gland 5

Post-pubertal mammary gland homeostasis 11

Pregnancy and lactational development of the mouse mammary gland 12

1.2. The case for mammary stem cells *13*

Historical perspective on mammary stem cell research: transplantation assays 13

New techniques for studying mammary stem cells: fate mapping in lineage tracing 15

Mammary stem cells: current perspectives 18

Multiplicity of mammary stem cells 18

A potential mammary stem cell niche 21

A novel approach for lineage tracing in the mammary gland 24

1.3. Imaging the murine mammary gland *26*

Historical approaches 26

A new technique for imaging the mammary gland 30

1.4. Aims of this thesis *31*

1.5. Importance and implications *32*

Chapter 2

Materials and methods **33**

2.1 Mice *34*

2.1.1 Husbandry 34

2.1.2 Mouse models 34

2.1.3 Initiation of lineage tracing 35

2.1.4 Tumours 35

2.1.5 EdU incorporation studies 35

2.2 Histology *36*
2.2.1 SeeDB optical clearing 36
2.2.2 CUBIC optical clearing 36
2.2.3 PACT-sRIMS optical clearing 37
2.2.4 Optical clearing and measurement of sample size changes 37
2.2.5 Human tumour optical clearing 37
2.2.6 Antibodies 38
2.2.7 EdU detection 38
2.2.8 Methyl green staining 39
2.2.9 Wholemount DAB staining 39
2.2.10 Organoid wholemount imaging 40

2.3 Mammary organoids *40*

2.4 Microscopy *41*
2.4.1 Confocal 41
2.4.2 Light sheet 42
2.4.3 Stereoscope 43

2.5 Imaging analysis *43*
2.5.1 Imaris 43
2.5.2 Computational analysis 43
2.5.3 PR / K8 association analysis 45
2.5.4 Nearest neighbour clonal analysis 45

2.6 Statistics. *46*

Chapter 3

Development of optical clearing and 3D imaging techniques for use in the mammary gland 47

3.1 Introduction *48*
3.1.1 Imaging requirements in the mammary gland 48
3.1.2 Optical clearing: theory and development 48
3.1.3 Aims 50

3.2 Results *52*
3.2.1 Assessment and selection of optical clearing protocols 52
3.2.1.1 SeeDB optical clearing 52
3.2.1.2 CUBIC optical clearing 55
3.2.1.3 PACT optical clearing 60
3.2.1.4 3DISCO optical clearing 64
3.2.2 Use of optical tissue clearing in cancer research 66
3.2.2.1 Optical clearing in mouse syngeneic tumour models 66
3.2.2.2 Optical clearing in human tumour samples and ex-vivo models 76
3.2.3 Combing optical clearing with other microscopy methods 78
3.2.3.1 Two-photon excitation microscopy 78

3.2.3.2 Light sheet fluorescence microscopy 80
3.2.4 Other applications of optical clearing methodologies 84
3.2.4.1 Liver 84
3.3 Discussion 86
3.3.1 SeeDB optical clearing 86
3.3.2 CUBIC optical clearing 87
3.3.3 Other clearing methods not tested 88
3.3.4 Other microscopy methods not tested 89
3.3.5 Conclusions 90

Chapter 4

Neutral lineage tracing in the mammary gland for studying development, homeostasis and stem cells. 92
4.1 Introduction 93
4.1.1 Mammary gland development 93
4.1.2 Lineage tracing in the mammary gland 94
4.1.3 Role of mammary stem cells 95
4.1.4 Aims of this chapter 96
4.1.5 Importance 97
4.2 Results 99
4.2.1 Tracing embryonic mammary stem cells 99
4.2.1.1 R26$^{[CA]30}$ SYNbglA 99
4.2.1.2 R26-Confetti 106
4.2.2 Tracing pubertal mammary stem cells 114
4.2.2.1 R26 $^{[CA]30}$ SYNgbIA 114
4.2.2.2 R26$^{[CA]30}$ YFP 119
4.2.2.3 R26-Confetti 132
4.2.3 Mammary stem cell tracing beyond pubertal morphogenesis 139
4.2.4 Tracing alveolar mammary stem cells 142
4.2.4.1 R26$^{[CA]30}$ YFP 142
4.2.4.2 R26-Confetti lineage tracing in the lactational mammary gland 149
4.2.4.3 R26$^{[CA]30}$ SYNbglA 154
4.3 Discussion 157
4.3.1 Summary 157
4.3.2 Overview of main findings 158
4.3.3 Conclusions 161

Chapter 5

Proliferation in the mammary gland 162
5.1 Introduction 163
5.1.1 Importance 164

5.1.2 Aims 164

5.2 Results *166*

5.2.1 Oestrous staging optimisation 166

5.2.2 Ki67 169

5.2.3 EdU 180

5.1.1.1 Pre-pubertal EdU proliferation 180

5.2.1.1 Pubertal proliferation 189

5.3.1.1 Post-pubertal EdU proliferation 195

5.2.4 Use of the Fucci2a system to study mammary gland proliferation 199

5.3 Discussion *204*

5.3.1 Overview of main findings 204

5.3.2 Oestrous cycling relevance 206

5.3.3 Importance and implications 207

5.3.4 Overall conclusions 207

Chapter 6

Assessment of an *ex vivo* organoid model to study mammary gland development 208

6.1 Introduction *209*

6.1.1 Background of 3D culture models in the mammary gland 209

6.1.2 Importance 211

6.1.3 Aims 211

6.2 Results *212*

6.2.1 Assessment of organoid biology 212

6.2.2 Investigating stromal cell presence in organoid cultures 223

6.2.3 Investigating proliferation in organoid cultures 228

6.2.4 Lineage tracing 234

6.3 Discussion *240*

6.3.1 Overview 240

6.3.2 Importance and implications 242

Chapter 7

Discussion 243

7.1 Context *244*

7.2 Defining a transcriptional profile of MaSCs *246*

7.2.1 Tracing of embryonic mammary stem cells 247

7.2.2 Examination of proliferation. 247

7.3 Overall conclusions *248*

Appendix 250

Genotyping protocols *251*

Confetti genotyping 251
CA30 genotyping 252

CUBIC reagent recipes *253*

Mammary organoid media *254*
Basal media 254
Complete media 254

Publications arising from this work *255*

References **305**

Table of figures

Figure 1.1 Schematic overview of murine embryonic mammary gland development 4
Figure 1.2 Schematic overview of murine mammary gland development 10
Figure 1.3 Limitations of 2D imaging in the mammary gland 29
Figure 3.2.1.1 SeeDB-based optical tissue clearing in the mammary gland 54
Figure 3.2.1.2 CUBIC-based optical tissue clearing in the mammary gland 56
Figure 3.2.1.1 CUBIC R1a-based optical tissue clearing in the mammary gland 58
Figure 3.2.1.2 CUBIC-based optical tissue clearing combined with colourimetric staining in the mammary gland 59
Figure 3.2.1.3 PACT-RapiClear optical tissue clearing in the mammary gland 62
Figure 3.2.1.4 PACT-RIMS and sRIMS optical tissue clearing in the mammary gland 63
Figure 3.2.1.5 3DISCO-based optical tissue clearing in the mammary gland 65
Figure 3.2.2.1 Assessment of optical tissue clearing protocols in mammary tumours 67
Figure 3.2.2.2 SeeDB optical clearing to examine apoptotic cell death in TUBO mammary tumours 68
Figure 3.2.2.3 Staining pattern in FFPE TUBO tumour sections 71
Figure 3.2.2.5 SeeDB optical clearing to examine cell proliferation and death in TUBO mammary tumours 73
Figure 3.2.2.6 SeeDB optical clearing to examine immune infiltrate in TUBO mammary tumours 75
Figure 3.2.2.7 CUBIC optical clearing in human tumours and collagen co-culture models 77
Figure 3.2.3.1 SeeDB optical clearing and 2-Photon excitation microscopy in murine TUBO tumours 79
Figure 3.2.3.2 Schematic diagram of light sheet microscopy imaging set up 82
Figure 3.2.3.3 CUBIC optical clearing and light sheet fluorescence microscopy in the mammary gland 83
Figure 3.2.4.1 CUBIC-based optical clearing in murine liver tissue 85
Figure 4.2.1.1 R26$^{[CA]30}$ SYNbglA tracing in the embryonic mammary gland 102
Figure 4.2.1.2 Clonal embryonic labelling in the R26$^{[CA]30}$ SYNbglA model 103
Figure 4.2.1.3. Clonal embryonic labelling in the R26$^{[CA]30}$ SYNbglA model 104
Figure 4.2.1.4 Immunohistochemistry of R26$^{[CA]30}$ SYNbglA clone shown in Figure 4.2.1.3. 105
Figure 4.2.1.5 Overview of the R26-Confetti model for embryonic tracing of mammary stem cells 109
Figure 4.2.1.6 Wholemount fluorescence stereomicroscopy images of the mammary ductal network observed in 6 week old R26R-Confetti mice labelled in utero. 110
Figure 4.2.1.7 Examples of multi-colour regions in R26R-Confetti mice in mammary glands of 6 week old mice labelled in labelled in utero. 111
Figure 4.2.1.8 Examples of uni-colour labelled distal regions in mammary glands of 6 week old R26R-Confetti mice labelled in utero. 112
Figure 4.2.1.9 Labelled distal regions in mammary glands of 6 week old R26R-Confetti mice labelled in utero. 113
Figure 4.2.2.1 Calculating the rate of strand slippage in lineage tracing of R26$^{[CA]30}$ SYNbglA pubertal animals. 116
Figure 4.2.2.2 Examples of large clonal regions in R26$^{[CA]30}$ SYNbglA pubertal animals. 117
Figure 4.2.2.3 Further examples of large clonal regions in R26$^{[CA]30}$ SYNbglA pubertal animals. 118
Figure 4.2.2.4 Overview of the R26$^{[CA]30}$ YFP mouse model for pubertal lineage tracing 123
Figure 4.2.2.5 Example of clonal labelling in R26$^{[CA]30}$ YFP mouse model during pubertal lineage tracing 124

Figure 4.2.2.6 Further example of clonal labelling in R26$^{[CA]30}$ YFP mouse model during pubertal lineage tracing ___ *125*
Figure 4.2.2.7 Further views of labelling in R26$^{[CA]30}$ YFP mouse mammary gland shown in Figure 4.2.1.6 _____ *126*
Figure 4.2.2.8. Computational analysis of clonal labelling in pubertal R26$^{[CA]30}$ YFP mammary gland shown in Figure 4.2.1.6& 7 _____ *127*
Figure 4.2.2.9 Analysis of differential K8 expression in luminal cells in R26$^{[CA]30}$ YFP mice _____ *128*
*Figure 4.2. 2.10 Further analysis of differential K8 expression in luminal cells in R26$^{[CA]30}$ YFP mice*_____ *129*
Figure 4.2.2.11. Volumetric computational analysis of clonal labelling in pubertal R26$^{[CA]30}$ YFP mammary gland shown in Figure 4.2.1.6&7 _____ *130*
Figure 4.2.2.12. Basal clonal labelling in pubertal R26$^{[CA]30}$ YFP mammary gland _____ *131*
Figure 4.2.2.13. Overview of the use of the R26R-Confetti model for lineage tracing during branching morphogenesis 134
*Figure 4.2.2.14. Example of rare CFP labelling in R26R-Confetti mammary glands.*_____ *135*
Figure 4.2.2.15. Labelling in R26R-Confetti mammary glands during branching morphogenesis after 1mg tamoxifen induction _____ *136*
Figure 4.2.2.16. Labelling in R26R-Confetti mammary glands during branching morphogenesis after 0.5mg tamoxifen induction _____ *137*
Figure 4.2.2.17. Lineage analysis in R26R-Confetti mammary glands during branching morphogenesis after 0.5mg tamoxifen induction _____ *138*
Figure 4.2.3.1 Lineage tracing in R26R-Confetti mammary glands during adult homeostasis, imaged with SeeDB optical clearing _____ *140*
Figure 4.2.3.2 Lineage tracing in R26R-Confetti mammary glands during adult homeostasis, imaged with SeeDB optical clearing _____ *141*
Figure 4.2.4.1 Overview of labelling in the R26$^{[CA]30}$ YFP mouse model for lineage tracing during lactation _____ *144*
Figure 4.2.4.2 Examples of different labelling patterns in the R26$^{[CA]30}$ YFP mouse model for during lactation _____ *145*
Figure 4.2.4.3 Luminal clonal regions in the R26$^{[CA]30}$ YFP mouse model for during lactation _____ *146*
Figure 4.2.4.4 Differential K8 expression seen in luminal cells of lactating R26$^{[CA]30}$ YFP mice _____ *147*
Figure 4.2.4.5 Basal clonal regions in the R26$^{[CA]30}$ YFP mouse model for during lactation _____ *148*
Figure 4.2.4.6 Overview of lineage tracing in R26R-Confetti mammary glands during lactation _____ *150*
Figure 4.2.4.7 Lineage tracing in R26R-Confetti mammary glands during lactation _____ *151*
Figure 4.2.4.8 Example of clonal regions seen in lineage tracing of R26R-Confetti mammary glands during lactation 152
*Figure 4.2.4.9 Quantification of lineage tracing in R26R-Confetti mammary glands during lactation*_____ *153*
Figure 4.2.4.10. Use of R26$^{[CA]30}$ SYNbglA mice to study MaSCs involved in multiple gestation / lactation / involution cycles _____ *155*
Figure 4.2.4.11 Representative labelling in R26$^{[CA]30}$ SYNbglA mice after involved in multiple gestation / lactation / involution cycles _____ *156*
Figure 5.2.1 Outline of method for determining oestrous stage in mice using vaginal lavage _____ *168*
Figure 5.2.2 Ki67 staining controls in the murine intestine and mammary gland _____ *171*
Figure 5.2.3 Ki67 staining in 4 week mammary glands _____ *172*
*Figure 5.2.4 Ki67 staining seen in ductal regions of 4 week old mammary glands*_____ *173*
Figure 5.2.5 Ki67 staining in terminal end ducts of 7 week mammary glands _____ *174*
Figure 5.2.6 Ki67 staining in 7 week mammary glands _____ *175*
Figure 5.2.7 Ki67 staining in 12 week mammary glands _____ *176*

Figure 5.2.8 Ki67 staining examples in 12 week mammary glands ____ *177*
Figure 5.2.9 Ki67 staining in 24 week mammary glands ____ *178*
Figure 5.2.10 Ki67 staining in 24 week mammary glands ____ *179*
Figure 5.2.11 EdU proliferation seen in small intestine control tissues ____ *182*
Figure 5.2.12 EdU proliferation in 4 week mammary glands terminal end buds ____ *183*
Figure 5.2.13 EdU incorportation in a bifucated TEB of 4 week mammary gland ____ *184*
Figure 5.2.14 EdU incorporation proximal to the nipple in 4 week mammary glands ____ *185*
Figure 5.2.15 EdU incorporation seen in the ductal network of 4 week mammary glands. ____ *186*
Figure 5.2.16 EdU incorporation seen in the side buds of 4 week mammary glands. ____ *187*
Figure 5.2.17 Population of K8 cells seen in 4 week pubertal mammary glands ____ *188*
Figure 5.2.18 EdU proliferation seen in TEBs of 7 week mammary glands ____ *190*
Figure 5.2.19 Absence of proliferation in 7 week mammary glands ____ *191*
Figure 5.2.20 Proliferation seen in stromal cells in 7 week mammary glands ____ *192*
Figure 5.2.21 Proliferation in ductal regions of 7 week mammary glands seen by EdU incorporation ____ *193*
Figure 5.2.22 Absence of proliferation in ductal regions of 7 week mammary glands ____ *194*
Figure 5.2.23 Proliferation in 12 week adult mammary glands seen by EdU incorporation ____ *196*
Figure 5.2.24 EdU incorporation investiageted in mammary glands of 24 week mice ____ *197*
Figure 5.2.25 Proliferation in 24 week mammary glands using EdU incorporation ____ *198*
Figure 5.2.26 Use of Fucci2a systm to study cell cycle in 12 week mammary glands ____ *201*
Figure 5.2.27 Fucci2a system in 12 week mammary glands shows cells in G1/S transition ____ *202*
Figure 5.2.28 Fucci2a system in 12 week mammary glands shows cells in S/G2/M transition ____ *203*
Figure 6.2.1 Time course showing the initial growth of mammary gland organoids ____ *215*
Figure 6.2.2 Staining controls for confocal microscopy of organoids ____ *216*
Figure 6.2.3 Basal lineage marker staining in mammary gland organoids: p63 ____ *217*
Figure 6.2.4 Basal lineage marker staining in mammary gland organoids: SMA and K5 ____ *218*
Figure 6.2.5 Luminal lineage marker staining in mammary gland organoids: K8 and E-cad ____ *219*
Figure 6.2.6 Luminal lineage marker staining in mammary gland organoids: hormone receptors ____ *220*
Figure 6.2.7 Mammary gland organoid formation in basal media ____ *221*
Figure 6.2.8 Lumen formation in mammary gland organoids using ZO1 ____ *222*
Figure 6.2.9 Example of immune cells in the in vivo mammary gland using CD45 ____ *224*
Figure 6.2.10 Brightfield images of stromal cells in mammary organoid cultures ____ *225*
Figure 6.2.11 CD45 cells in mammary gland organoid cultures ____ *226*
Figure 6.2.12 CD45 cells in mammary gland organoid cultures ____ *227*
Figure 6.2.13 Examination of proliferation in mammary organoids using Ki67 staining ____ *229*
Figure 6.2.14 EdU staining control in mammary organoids ____ *230*
Figure 6.2.15 Examination of proliferation in mammary organoids using EdU incorporation ____ *231*
Figure 6.2.16 Examination of proliferation in mammary organoids using EdU incorporation ____ *232*
Figure 6.2.17 Examination of proliferation in mammary organoids using EdU incorporation ____ *233*
Figure 6.2.18 Examples of small clonal regions in $R26^{[CA]30}$ YFP mammary organoids ____ *236*
Figure 6.2.19 Examples of medium-sized clonal regions in $R26^{[CA]30}$ YFP mammary organoids ____ *237*
Figure 6.2.20 Example of a large clonal region in $R26^{[CA]30}$ YFP mammary organoids ____ *238*

Figure 6.2.2I Example of a large clonal region in R26$^{[CA]30}$ YFP mammary organoids _____ *239*
Figure 7.1 Schematic for proposed model of mammary stem cells _____ *246*
Figure 7.2 Working model of the mammary epithelial cell hierarchy _____ *249*

Abbreviations

2D	Two dimensions
3D	Three dimensions
3DISCO	3D imaging of solvent-cleared organs
BRDU	Bromodeoxyuridine
BSA	Bovine serum albumin
CARS	Coherent anti-stokes Raman spectroscopy
CC3	Cleaved caspase 3
CFP	Cyan fluorescent protein
CSF-1	Colony stimulating factor-1
CUBIC	Clear unobstructed brain imaging cocktails
DAB	Diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
DBE	Dibenzyl ether
DCM	Dichloromethane
DLL1	Delta-like 1
E	Embryonic day
E-CAD	E-cadherin
EDU	5-ethynyl-2'-deoxyuridine
ER	Oestrogen receptor
FACS	Fluorescence-activated cell sorting
FFPE	Formalin-fixed paraffin-embedded
FUCCI	Fluorescent ubiquitination-based cell cycle indicator
GFP	Green fluorescent protein

HCL	Hydrogen chloride
HMSC	Human mesenchymal stem cells
HRP	Horseradish peroxidase
I.P	Intraperitoneal
IGF1	Insulin-like growth factor 1
K14	Cytokeratin 14
K8	Cytokeratin 8
LAS	Leica acquisition software
LSFM	Light sheet fluorescence microscopy
MASC	Mammary stem cells
MMP	Matrix metalloproteinases
MRU	Mammary repopulating unit
NBF	Neutral buffered formalin
NGF	Nerve growth factor
NGS	Normal goat serum
NM	Nanometres
PACT	Passive CLARITY Technique
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
PI-MEC	Parity-identified mammary epithelial cell
PR	Progesterone receptor
PROCR	Protein C receptor
R1	(CUBIC) reagent 1
R2	(CUBIC) reagent 2
R26	Rosa26

RFP	Red fluorescent protein
RI	Refractive index
RIMS	Refractive index matching solution
SD	Standard deviation
SEEDB	See deep brain
SHG	Second harmonic generation
SMA	Smooth muscle actin
SYNBGLA	Beta-glucosidase
TEB	Terminal end bud
THF	Tetrahydrofuran
UDISCO	Ultimate 3D imaging of solvent-cleared organs
VEGF	Vascular endothelial growth factor
WAP	Whey acidic protein
YFP	Yellow fluorescent protein

Chapter 1

Introduction

1.1. Introduction to the mammary gland

The mammary gland is a complex organ, which undergoes multiple cycles of expansion and regression throughout life. Unlike most organs, the majority of mammary gland development occurs postnatally. During puberty there is a large expansion of the epithelial compartment, with elongation and branching of the ductal tree throughout the mammary fat pad. The mammary gland also undergoes significant growth and tertiary branching during pregnancy to form milk-producing alveoli that are removed during post-lactational regression by rapid cell death coupled with tissue remodelling. As such, the mammary gland represents a unique organ to study adult stem cells and their contribution to tissue development, maintenance and remodelling.

Moreover, as one of the defining features of mammalian biology, study of the mammary gland has important evolutionary interest. Indeed, although several aspects of mammary gland biology vary between mammals – for example number and positioning – development of the mammary gland amongst mammals broadly follows the same stages (Oftedal and Dhouailly, 2013). Consequently, the mouse is widely used to model mammary developmental mechanisms due to the functional similarities between the mammary tissue of both human and mouse, in addition to its amenability to complex genetic manipulations.

Embryonic mammary development of the mouse

During mouse embryonic development, mammary placode formation is initiated at approximately embryonic day (E) **10.5,** with the formation of two so-called milk lines from the overlying ectoderm (Propper et al., 2013). These mammary lines, which can be detected by expression of a number of Wnt proteins including Wnt10b (Veltmaat et al., 2004) and Wnt5a (Yamaguchi et al., 1999), form bilaterally along the rostral/caudal axis between the fore and hind limb buds (Sakakura et al., 1982). This is followed a day later on **E11.5** with the development of five pairs of placodes along these lines (Figure 1.2,i), formed of aggregates of epithelial cells which can be detected histologically by the thickening of the surface ectoderm. Interestingly, these placodes develop asynchronously in a defined order, but with each individual pair of placodes developing symmetrically (Veltmaat et al., 2004, 2006). By **E12.5,** the cells of the placode expand then begin to pile up and arrange themselves into a concentric formation, meaning they can be detected as small bulges on the ventral surface of the embryo (Howard, 2012) (Figure 1.1,ii).

By **E13.5**, each placode expands and invaginates into the underlying dermal mesenchyme (Figure 1.1,iii), to result in the formation of a mammary bud which continues to expand until **E14.5** (Watson and Khaled, 2008). At this point, the underlying mesenchyme also begins to change as the surrounding cells elongate and condense around the expanding mammary bud (Figure 1.1,iv). In male mice, this time coincides with the production of androgens leading to the expansion of the underlying mesenchyme, which eventually severs the connection of the mammary bud to the skin (Heuberger et al., 1982; Veltmaat et al., 2003) and is followed by apoptosis of the mammary mesenchyme and the majority of the mammary bud (Dunbar et al., 1999). This process is in contrast to male human development, whereby this severing event does not occur. Instead, the bud goes on to form a small ductal outgrowth that remains connected to the nipple (Howard and Gusterson, 2000). For this reason, although rare, male humans can be susceptible to breast cancer, in addition to other benign aberrations of the breast.

In female mice, by **E15.5** distinct cell types form within the neck of the mammary bud which subsequently push the distal end of the mammary bud deeper (Figure 1.1,v). Consequently, **E16.5** marks the commencement of the first phase of branching morphogenesis in the mammary gland. Here, the mammary bud invaginates further from the primary mammary mesenchyme and into a third stromal compartment (Figure 1.1,vi). Termed the mammary fat pad pre-cursor, this consists of an assemblage of preadipocytes which mediate a number of paracrine signalling pathways (Hens and Wysolmerski, 2005). The invasion into the mammary fat pad initiates the expansion and branching of the epithelial tubular sprout, culminating in the development of the primordial ductal tree by **E18.5** (Figure 1.1,vii) (Howard, 2012). From this point onwards, the embryonic mammary gland is unusual in that it remains relatively quiescent, expanding only isometrically with the rest of the body, even after birth. At the termination of embryonic mammary development, the epithelium of the gland consists of a primary duct consisting of approximately 10 – 15 initial branches surrounded by a number of stromal cells, altogether which forms the scaffold upon which further mammary development will be built.

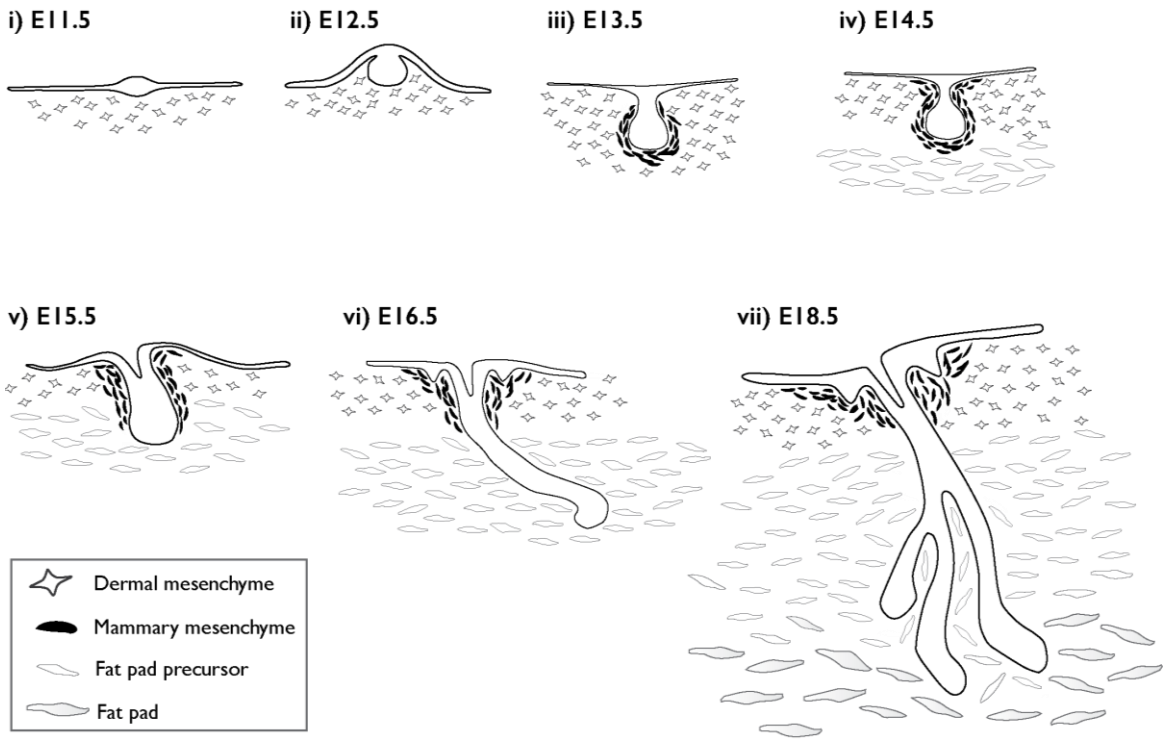


Figure 1.1 Schematic overview of murine embryonic mammary gland development

(i) Mammary placodes form at E11.5 along the milk lines, visible as a slight thickening of the endoderm. **(ii)** At E12.5, the mammary placode can be detected as visible elevation as the mammary bud forms. **(iii)** At E13.5 the mammary mesenchyme expands as the placode invaginates into the underlying dermal mesenchyme. **(iv)** At E14.5, further expansion can be seen as the mammary mesenchyme thickens further and the fat pad precursor is formed. **(v)** At E15.5 the mammary bud elongates further and reaches the fat pad precursor which initiates a round of branching morphogenesis **(iv)** at E16.5. By E18.5 **(vii)** a primordial ductal tree is present which expands isometrically until puberty. Based upon figure from Veltmaat *et al*., 2003.

Pubertal development of the mouse mammary gland

Interestingly, despite being just a rudimentary ductal system, the gland at birth is equipped with the capability to produce milk, colloquially referred to as witch's milk in humans. This is due to foetal exposure to maternal hormones, either from the placenta during the latter stages of gestation or via the breast milk from the mother herself after birth. However, the hormones responsible for this process do not cause significant further expansion of the ductal network, and once this aberrant endocrine influence has subsided the gland returns to a period of isometric growth, expanding only at a rate proportionate with the overall growth of the rest of the body.

In the mouse, the symphony of hormone and growth factor signalling that initiates puberty begins at approximately 4 weeks of age (Elo et al., 2017), where the growth of the gland switches to allometric growth. Here, the gland returns to a highly dynamic state, largely as a response to the increase in serum levels of oestrogen, a membrane-soluble ligand produced by the ovaries, which acts on its two nuclear receptors, ERα and ERβ, to generate the burst of proliferation needed to re-enter another stage of branching morphogenesis. In concert with this, proliferation is also mediated through paracrine and endocrine insulin-like growth factor 1 (IGF1) signalling in the mammary stroma, which is produced both locally in the mammary gland and systemically in the liver via pituitary growth hormone action (Ruan and Kleinberg, 1999; Richards et al., 2004). The establishment of the requirement for oestrogen in pubertal morphogenesis has been shown by a variety of knockout mouse models, which also demonstrated that it is ERα that is the most important for development, with complete knockout of the receptor resulting in the inability of ducts to invade the fat pad (Mallepell et al., 2006; Feng et al., 2007).

This proliferation and invasion of the fat pad is driven by the formation of terminal end buds (TEBs) at the tip of ducts (Figure 1.2, iv). These club-shaped structures are comprised of an outer compartment formed of a single layer of cap cells and a multi-layered inner core of body cells, which are found 4 – 6 cells thick (Williams and Daniel, 1983). As the TEBs continue to penetrate the fat pad, clefting in the highly proliferative tip of the structure occurs which results in the subsequent bifurcation (and even trifurcation) to generate the branches that form the ductal tree (Figure 1.2,ii). This is thought to be primarily lead by the cap cells, which have been shown to have a superior ability to regenerate the ductal tree in in transplantation

studies (Bai and Rohrschneider, 2010) and so have historically been considered as a reservoir of regenerative mammary stem cells (MaSCs).

The cells of these invading TEBs subsequently give rise to the expanding mammary epithelium, with the more differentiated and least proliferative cells found in the neck and subtending duct. As the duct elongates the TEB cells differentiate into the two main cell lineages of the mammary epithelium: luminal epithelial cells, marked by keratin (K)6, 8 and 18, are presumed to form from body cells and line the central lumen of the ductal structures. It is a subset of the luminal cell compartment which subsequently differentiates into secretory cells to produce milk during pregnancy and lactation, and as such express oestrogen and progesterone receptors to allow them to respond to hormonal signals (Ismail et al., 2002; Grimm et al., 2006). The cap cell layer is thought to differentiate into the other epithelial cell type of the mammary gland that makes up the outer layer of the duct, the myoepithelial cells (herein referred to as basal cells) (Tiede and Kang, 2011). These are the cells that are in contact with the surrounding basement membrane.

The interaction between cap and body cells of the TEB is a subject of dynamic research in and of itself: initially, it was hypothesised that the outer layer of cap cells are able to enter the lumen of the TEB to contribute to the body cells, and thus the subsequent luminal cell lineage. This was supported by later work, which demonstrated that body cells within the TEB undergo apoptosis (Humphreys et al., 1996), potentially to support lumen formation in the duct. However, more recent work utilising mathematical modelling has demonstrated that cap cells do not contribute to the body and luminal cell lineage (Paine et al., 2016). Nevertheless, recent innovative *in vivo* imaging of the TEB has shown for the first time, that the TEB is a highly dynamic structure within itself, with cell migration and movement occurring constantly (Scheele et al., 2017). Altogether, this shows the unique nature of TEBs – given their heterogeneous cellular composition, high proliferation rates (60–90%) at the same time as high apoptosis rates (5–15%), level of invasive ability, and the ability to recruit stromal cells (Daniel and Smith, 1999).

However, ductal elongation is not only an epithelial intrinsic mechanism. Multiple studies have demonstrated the instructive role of the mammary stroma for ductal morphogenesis, and its fundamental role in dictating proper ductal development. Accordingly, the mammary stroma is comprised of several cell types: the first and most abundant is the adipocytes, which form

the mammary fat pad that develops in concert with the epithelia during embryonic development. Adipogenic signalling in the mammary gland has been shown to be important (Couldrey et al., 2002); indeed, this is emphasised by the effect of the mammary bud reaching the pre-adipocytes in the fat pad pre-cursor during embryonic development, highlighting the tight association between adipocytes and mammary epithelial cells throughout development.

Several different immune cells are also a vital component of the mammary stroma. In particular, macrophages have been found to be in tight association with the mammary epithelium during pubertal development, even being detected within the body cell layer where they have been shown to support the formation of the lumen in the subtending duct by clearance of apoptotic cells via phagocytosis (Gouon-Evans et al., 2000). Macrophages also function within the stroma surrounding the TEBs, where they are recruited by colony stimulating factor-1 (CSF-1) which also regulates their survival and proliferation (Van Nguyen and Pollard, 2002). Here, they are found within the collagen fibres of the basement membrane where they are presumed to help guide the direction of the developing TEB. This function is thought to be due to their role in remodelling collagen I into long fibrillar bundles that project laterally from the sides of the TEB (Ingman et al., 2006). Recent work has developed this further, demonstrating that macrophage function in the mammary gland is STAT5 dependent (Brady et al., 2017).

In addition to contributing to the structural morphology and organisation of the gland, macrophages also function as one of the many paracrine signalling mediators in the mammary gland. Furthermore, macrophages have been shown to influence stem cell activity and number in the mammary gland, demonstrated by the reduction in mammary repopulating units in transplantation studies (Gyorki et al., 2009). Until recently the exact mechanism was not fully understood. However, recent work has suggested that this may be mediated via a feedback loop involving Notch and Wnt signalling pathways, which are well known to be involved in adult stem cell maintenance, including in the mammary gland (Chu et al., 2004; Plaks et al., 2013; Lilja et al., 2018). It is proposed that interaction occurs within a "macrophageal niche", which is identified exclusively in the mammary gland. Here, a subpopulation of MaSCs express the Notch ligand Delta-like 1 (Dll1), which activates Notch2/3 receptors on macrophages within the mammary stroma to cause downstream Notch signalling. This subsequently induces expression of Wnt ligands which initiates a feedback loop, whereby the Wnt ligands further promote the function of these Dll1+ MaSCs (Chakrabarti et al., 2018).

Other more surprising immune cell populations contribute to the milieu of factors that make up the mammary stroma. Granulocytes, which are traditionally considered to be primarily involved in allergic responses, pathogen defence and wound healing can be found within the pubertal mammary stroma. Recruitment of eosinophils is mediated by secretion of the chemokines eotaxin and amphiregulin, with the latter being induced by oestrogen and progesterone production (Aupperlee et al., 2014). These cells can be found surrounding the leading tip of the TEB where they are thought to be involved in branch patterning – indeed, the absence of eosinophils leads to a reduction in branch number and TEB formation, although this does not affect ductal elongation (Gouon-Evans et al., 2000). The precise mechanism for this is not fully understood, but eosinophils themselves have been shown in other systems to produce several cytokines and growth factors which have been shown to be important for ductal morphogenesis, such as vascular endothelial growth factor (VEGF) (Horiuchi and Weller, 1997). Moreover, eosinophils also appear to secrete chemokines in order to mediate the recruitment and growth of other immune cells, including macrophages via the chemoattractant C10 (Gouon-Evans et al., 2002), and mast cells via eosinophil-derived stem cell factor and nerve growth factor (NGF), which is crucial for mast cell survival and activation (Rothenberg and Hogan, 2006).

Mast cells themselves, another type of granulocyte, have also been shown to be important in the regulation of proliferation in TEBs and subsequent ductal elongation, with aberrations in both number and function of mast cells affecting pubertal mammary development (Reed and Schwertfeger, 2010). In particular, full degranulation of mast cells has been highlighted as an essential aspect. The release of active intracellular factors has been shown to be crucial for mast cells mediating their effects, such as serine proteases, which may act on the extracellular matrix to remodel the stroma. The actual amount of mast cells in the gland is also critical: in the *W-sash* genetic mouse model for mast cell ablation, ductal tree development was more affected in mice homozygous for *W-sash* relative to their heterozygote and wildtype littermates (Lilla and Werb, 2010).

The vital significance of the mammary stroma throughout mammary development and beyond has been elegantly highlighted by the transplantation of mammary epithelium into the mesenchyme of various other tissues, including that of the salivary gland (Kratochwil, 1969; Sakakura et al., 1976), prostate (Taylor et al., 2009) and epidermal lineages (Cunha et al., 1995),

whereby mammary epithelium is able to contribute to, and form the morphology inherent to the transplanted environment. However, memory of mammary cell identity can still be maintained, as shown by the production of milk proteins following stimulation with pregnancy hormones after mammary cells have been transplanted into the salivary mesenchyme (Sakakura et al., 1976).

Conversely, non-mammary cells can be transplanted into the mammary stroma and will adopt a mammary cell fate – as has been demonstrated for several cell types including normal, stem/progenitor (Booth et al., 2008; Boulanger et al., 2012), and cancer cells (Booth et al., 2011). Indeed, even testicular cells can alter their cell fate and contribute towards the regeneration of the ductal tree (Boulanger et al., 2007). Of note, there is a requirement for co-transplantation of mammary epithelial cells, suggesting a potential role for paracrine signalling of mammary cells in the transdifferentiation of cell fate – however whether this is mediated via direct signalling to the foreign cells or rather indirectly, for example by activating the naive stroma, remains to be determined (Howard and Lu, 2014). Together, these transplantation studies establish that mammary epithelial cell identity appears to be dictated by a combination of both local cues from the mammary stroma and cell-intrinsic signals.

In the mouse, TEBs regress at approximately 10 weeks of age when the ductal network has reached the extremities of the fat pad, leaving behind blunt ended ductal termini (Paine and Lewis, 2017). Thus, pubertal development and ductal morphogenesis of the mammary gland is ceased.

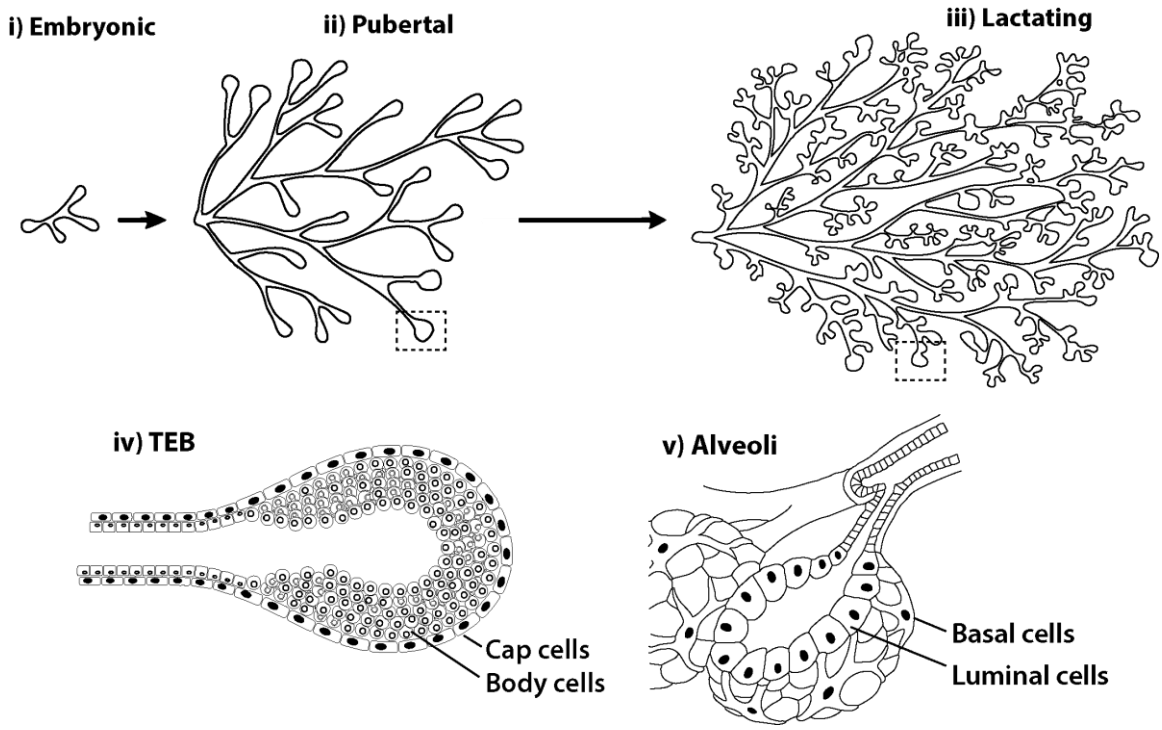


Figure 1.2 Schematic overview of murine mammary gland development

i) A rudimentary ductal tree develops during embryogenesis of the mammary gland, which expands during puberty with the onset of ovarian hormone production **(ii)**. During this time, terminal end buds (TEB) form **(iv)**, composed of cap and body cells. After the fat pad is filled at the end of puberty, the gland is also able to undergo further expansion with lactation **(iii)**, whereby lateral side branching occurs with the formation of lobuloalveolar structures **(v)**, formed again of luminal and basal cells. Figure based on Gjorevski and Nelson, 2011.

Post-pubertal mammary gland homeostasis

After pubertal branching morphogenesis is complete the post-pubescent mammary gland is often said to remain relatively dormant until pregnancy. However, we are gaining increasing understanding in both humans and rodents, that this may be in fact a misnomer. Indeed, in mice morphological and cellular changes occur over the short 4-5 day oestrous cycle, with the mammary gland exhibiting a pattern of mild proliferation, differentiation and involution (Cole, 1933). During this period, lateral expansion results in the formation of tertiary branches, and brief pulses of alveolar budding occurs cyclically in response to ovarian hormones, with the gland preparing for pregnancy, which then collapse in response to hormonal cues if pregnancy does not occur at the end of the oestrous cycle.

The mouse oestrous cycle can be split into four parts, largely defined by ovarian hormone levels: ovulation occurs during the oestrous stage of the cycle, which can be characterised by maximum serum oestrogen levels (Zeps et al., 1999). Conversely, a substantial drop in progesterone levels occurs due to the regression of the corpus luteum if implantation does not occur (Fata et al., 2007). Accordingly, the mammary epithelium is largely of ductal morphology and it is during this stage that that gland is relatively quiescent, with minimal levels of both proliferation and apoptosis detected.

Following this, metoestrous occurs, as serum levels of progesterone begin to rise, with oestrogen levels dropping. This is also associated with increases in expression of matrix metalloproteinases (MMPs) within the mammary gland, including MMP-9 and MMP-13 (Fata et al., 2003) which are presumed to mediate oestrous-cycle related remodelling ready for expansion of the ducts (Khokha and Werb, 2011). Accordingly, macrophages also begin to increase in both abundance and association with the mammary epithelium where they presumably aid in this remodelling (Hodson et al., 2013).

Dioestrous, characterised by peak serum levels of progesterone and reduction in oestrogen, is associated with the process of tertiary branching and lobuloalveologenesis. The resulting increase in mammary epithelial volume is accompanied by peak cell proliferation and followed by cell death at the end (Giraddi et al., 2015). As such, dioestrous has sometimes been divided into two, which has been recently described further whereby two dioestrous states exist, with either low or high proliferative activity, as shown by Ki67 labelling (Shehata et al., 2018). This parallels the luteal stage in the human menstrual cycle (Navarrete et al., 2005), where

increased density in mammograms has been demonstrated (Ursin et al., 1998). These lobuloalveolar buds formed in the mouse are rudimentary alveolar structures which are only capable of milk production when pregnancy-induced growth, mediated by prolactin and progesterone, fully differentiates them into their full grown secretory lobuloalveolar structures.

The final stage is pro-oestrous, whereby collapsing of the alveolar epithelium takes place if pregnancy does not occur. This process is mediated by the sharp decline in serum levels of progesterone, as oestrogen levels begin to return (Fata et al., 2001). Macrophages again play a role, where they are thought to help mediate this remodelling and cell death (Chua et al., 2010).

Pregnancy and lactational development of the mouse mammary gland

In the event of pregnancy, the mammary gland completely changes form once again, whereby multiple signalling pathways are activated to orchestrate an unprecedented amount of proliferation resulting in alveologenesis. Specifically, increases in the levels of progesterone stimulate the proliferation and initiate the formation of tertiary side branches and the secretory alveoli responsible for the production of milk in response to prolactin (Gjorevski and Nelson, 2011). The mammary stroma is also remodelled, with a reduction in volume of the adipocyte compartment to accommodate for the vast expansion of the epithelium. Furthermore, the vasculature is remodelled such that capillaries can be found in close contact with each individual alveolus. Subsequent cleavage and differentiation of the alveolar buds occurs in the second stage of pregnancy resulting in the formation of individual alveoli.

Following birth, oxytocin is induced by suckling during lactation, which stimulates the contraction of alveolar basal cells to expel milk from the lumen of the alveoli. This dynamic process continues throughout lactation until weaning, when prolactin levels decline and milk production ceases (Hynes and Watson, 2010). Post-lactational regression of the mammary gland, known as involution, then occurs, marked by extensive programmed cell death of secretory alveoli and the re-emergence of adipocytes (Kreuzaler et al., 2011). Interestingly, this first phase of involution is reversible, with the gland able to re-initiate lactation upon further suckling. However, the second phase is irreversible - whereby degradation of the basement membrane occurs by MMP2 and MMP3 expressed by the mesenchyme (Fata et al., 2007). Consequently, the mammary gland is returned to a structure akin to the nulliparous

state formed of a network of ductal branches (Sargeant et al., 2014). As such, this cycle of alveologenesis will occur anew with each successive pregnancy.

1.2. The case for mammary stem cells

This remarkable capacity of the mammary gland for rapid growth and regeneration through multiple cycles of both oestrus and pregnancy has been attributed to the existence of adult mammary stem cells (MaSCs). Although the identification and characterisation of these putative stem cells is a highly dynamic area of research that has extended over 60 years, there remains significant controversy and disparity regarding the number, molecular identity, anatomical location and differentiation hierarchy of these elusive MaSCs.

The presence of adult stem cells has long been known in a variety of different epithelial tissues, such as the intestine, skin, liver and stomach (Tumbar et al., 2004; Lopez-Garcia et al., 2010; Stange et al., 2013; Huch et al., 2014). However, in these organs the key role of adult stem cells in to maintain homeostasis of the tissue – namely the replacement of cells which are lost due to attrition or injury. In contrast, the majority of mammary gland development occurs postnatally, as outlined previously. Consequently, the adult stem cells within the mammary gland are required to serve both developmental and homeostatic purposes, in addition to allowing the gland to completely change form in pregnancy. To study these important cells, two main approaches have been historically applied, which will be discussed below.

Historical perspective on mammary stem cell research: transplantation assays

Transplantation studies carried out in the 1950's was the first methodology used that provided evidence of the existence of MaSCs. This seminal work described the use of de-epithelialised ('cleared') mouse mammary fat pads, whereby endogenous epithelium is surgically removed, as transplantation sites for normal, pre-neoplastic or malignant mammary tissue fragments (DeOme et al., 1959). Portions of normal mammary epithelium could be successfully engrafted into cleared fat pads of recipient mice, resulting in the regeneration of an entirely new ductal epithelial tree *in vivo*. This was the first suggestion of the existence of repopulating cells in the mammary gland, inferred to be MaSCs (also called mammary repopulating units, MRUs) (Hoshino, 1962; Smith and Medina, 1988). Subsequently, studies showed successful engraftment from any segment of the mammary epithelial tree, indicating the widespread distribution of MaSCs (Smith and Medina, 1988). Notably, these outgrowths could in turn

produce secondary outgrowths when serially transplanted, confirming that mammary epithelium contained cells that displayed characteristics of stem cells, such as self-renewing potential and multipotency (Oakes et al., 2014).

Subsequently, using fluorescence-activated cell sorting (FACS) it was shown that that a single cell had the capacity to reconstitute the entire mammary epithelium. The first study to suggest this relied on experiments using retrovirally marked mammary epithelial cells as endogenous clonal markers (Kordon and Smith, 1998). Subsequent work relied on the differential expression of cell surface markers to identify and isolate heterogeneous mammary epithelial cell subpopulations in order to provide functional evidence that a single multipotent cell could regenerate the entire mammary epithelium upon transplantation (Shackleton et al., 2006; Sleeman et al., 2006; Stingl et al., 2006). This was in turn further refined, showing that specific subpopulations of cells could be purified which had superior repopulating capacities.

As such, transplanting single cells and assessing their ability to clonally expand to form a functional mammary gland comprised of both cellular lineages, in addition to their ability to self-renew in serial transplants, has been traditionally considered the "Gold Standard" assay for the detection of MaSCs (Visvader and Stingl, 2014). In addition to suggesting that MaSCs were sporadically located throughout the gland, transplantation studies also inferred that they were present at various stages throughout development, and that they were long lived (Kordon and Smith, 1998). Moreover, together these analyses supported the notion that adult MaSCs are bi- or multipotent in nature, in that they are able to produce cells of both the luminal and basal lineage, in addition to the secretory luminal lineage during lactation.

However, recent lineage tracing studies have raised concerns regarding the reliability of transplantation assays for investigating the cell fate of normal mammary epithelial cells. Indeed, it has been subsequently demonstrated that lineage-restricted cells can be forced to adopt a multipotent fate under "regenerative conditions", which has challenged the previous dogma (Van Keymeulen et al., 2011; de Visser et al., 2012; van Amerongen et al., 2012). It is postulated that the process of tissue dissociation and transplantation into a cleared fat pad may reveal properties of cells that do not reflect their true behaviour in intact tissue *in vivo*. Transplantation assays are therefore increasingly considered more akin to an injury model, where cells may de-differentiate to reacquire stem-like properties, rather than a way of assaying for a native stem cell population (Visvader and Stingl, 2014).

Consequently, it is now largely accepted that the mammary repopulating cells that are identified by transplantation are distinct from the stem cells that exist under physiological conditions. Nevertheless, this technique has provided some important insights into the qualities of self-renewal and regeneration, with enduring relevance. Moreover, the notion that fate decisions within the hierarchy are not strictly unidirectional, and in some conditions can be reversed, has wide-reaching implications for both oncology and regenerative medicine.

New techniques for studying mammary stem cells: fate mapping in lineage tracing

While transplantation assays are useful for assessing the repopulating capacity of defined subpopulations, the advent of mouse models engineered to express reporter proteins that are driven by pre-specified, lineage-specific promoters has enabled the tracking and fate mapping of putative MaSCs *in vivo* under physiological conditions. These studies utilise tamoxifen- or doxycycline-responsive transgenic mouse models to induce the expression of reporter genes in predefined populations of cells (Van Keymeulen et al., 2011; de Visser et al., 2012; van Amerongen et al., 2012; Rios et al., 2014; Wang et al., 2015; Davis et al., 2016; Wuidart et al., 2016). The genetic label, which is typically either a fluorescent or histochemical reporter, is then permanently expressed by the original cell and can be subsequently transmitted to all of its progeny. An analysis of reporter expression through time can then be used to determine whether the original labelled population contained lineage-restricted stem cells or cells with multi-lineage differentiation potential.

As such, studies have employed various gene promoters in the mammary gland to trace the fate of different cell populations, including Elf5, a proposed marker of luminal progenitor cells (Oettgen et al., 1999; Oakes et al., 2008), and various cytokeratin's reported to be differentially expressed in the various cell lineages. The first inducible cell fate mapping study to use this system in the mammary gland challenged prevailing work on the contribution of multipotent MaSCs to postnatal gland development. This study used a transgenic inducible Cre recombinase driven by K14/K5 or K8/K18 promoters for labelling basal and luminal cells and their progeny, respectively. Inducing labelling of K14-expressing cells in embryos resulted in the labelling of both luminal and basal cells at puberty, implying that embryonic basal cells are multipotent.

However, when labelling was induced after birth unipotent cells were traced throughout the different stages of development, suggesting that lineage-restricted MaSCs, rather than bipotent stem cells, drive development and homeostasis in the postnatal mammary gland (Van Keymeulen et al., 2011). Moreover, when the differentiation potential of both epithelial lineages was determined using transplantation assays it was demonstrated that label-positive basal cells, but not luminal cells, were capable of reconstituting a mammary gland, in line with previous findings that originally identified multipotent basal MaSCs. In turn, this strongly supports that the experimental context of transplantation forces differentiation of basal MaSCs into both epithelial lineages, while in the intact gland *in vivo* only lineage-restricted unipotent stem cells exist.

Controversially, other studies using similar fate mapping methodologies then provided further evidence in support of the original paradigm for the existence of bipotent stem cells in the adult mammary gland. Here, bipotent MaSCs traced at a clonal level *in situ* demonstrated that basal-labelled cells generated both cellular lineages, were long lived, and contributed not only to the major stages of morphogenesis in the postnatal gland but also in ductal tree maintenance during adult homeostasis (Rios et al., 2014). Since this work, another study has also provided evidence for the existence of unique, multipotent MaSCs that are marked by the expression of protein C receptor (Procr), a novel Wnt signalling target in the mammary gland (Wang et al., 2015).

Subsequent lineage-tracing studies have provided evidence in support of both unipotent (Davis et al., 2016; Wuidart et al., 2016; Scheele et al., 2017) and bi/multipotent adult MaSCs (van Amerongen et al., 2012). Furthermore, lineage-restricted cell populations have been shown to possess the capacity to convert to multipotency *in vivo* by oncogenic PI3KCA signalling (Koren et al., 2015; Van Keymeulen et al., 2015) - suggesting that there is scope for plastic transformation and thereby adding further complexity to this system.

Consequently, these inconsistencies in the results gained from lineage tracing in the normal mammary gland have therefore raised questions regarding the accuracy of some of these methods and brought to light some of their drawbacks. Firstly, a key limitation of many of the fate mapping approaches is the reliance on promoter driven reporter expression. This leads to assumptions regarding the molecular identity of MaSCs and progenitors, which therefore lends itself to bias and limitations in the cells that can be traced. Moreover, the temporal

expression of pathway-specific promoters or the fidelity of pan-lineage promoters means previous models are now being re-examined (de Visser et al., 2012; van Amerongen et al., 2012). Accordingly, given that a single mammary stem/progenitor cell is capable of producing many hundred progeny (Davis et al., 2016) the promiscuous labelling of even a small number of cells of the opposing lineage could significantly confound downstream lineage analysis in this model (Wuidart et al., 2016). Secondly, criticisms have also arisen when high doses of tamoxifen are used to induce promoter driven reporter expression, given its effect on oestrogen signalling (Asselin-Labat et al., 2010; Shehata et al., 2014).

A third limitation relates to the power of population-based labelling approaches to accurately detect the expansion of a single clone, which is a function of both the method of detection and the initial labelling density. To overcome this problem, as well as potential tracing artefacts associated with the preferential labelling of specific (and potentially non-representative) cell subpopulations, a recent study has mapped the fate of all basal cells (a technique termed saturation lineage-tracing) (Wuidart et al., 2016). If rare bipotent MaSCs do reside in the basal compartment and contribute even minimally to mammary gland morphogenesis and homeostasis (Visvader and Clevers, 2016), this could be detected by an increase in the number of fluorescently labelled luminal cells observed using either FACS or 3D image quantification. However, no population flux was detected using either method of analysis in these studies, suggesting that basal MaSCs are indeed lineage-restricted.

In the developing gland, stem cells have been proposed to reside in the proliferating TEBs, particularly cap cells. These data raise important questions - do stem or long-lived progenitor cells remain in the mature mammary gland after ductal morphogenesis is complete? And, if so, where do these stem/progenitor cells reside after the TEBs collapse at the end of puberty? One possible explanation is that MaSCs may be deposited along the length of the ducts by the travelling TEBs during ductal elongation. Moreover, it may be that these proliferating MaSCs seen during puberty then become quiescent in adulthood when the requirement for extensive proliferation in the virgin gland ceases. Additionally, although these preliminary observations require further characterisation, elucidation of how these presumptive MaSCs are deposited along the length of the ductal tree (e.g., whether this occurs in a random fashion or during bifurcation of a TEB or upon side branching) is an important question, and may shed new light on the likelihood of a mammary stem cell niche.

Mammary stem cells: current perspectives

As a result of the work spanning 60 years, there are now two main schools of thought regarding the identity of MaSCs. The first, largely brought about through lineage tracing experiments, suggest that each of the main cell types is derived from their own unipotent, lineage-restricted stem cell (Van Keymeulen et al., 2011; Davis et al., 2016; Wuidart et al., 2016; Lloyd-Lewis et al., 2017; Scheele et al., 2017). The second hypothesis suggests that bipotent MaSCs exist that are able to give rise to both luminal and basal cell lineages (Shackleton et al., 2006; Rios et al., 2014; Wang et al., 2015); as supported by the ability of single cells to repopulate both lineages in a cleared mammary gland. While both transplantation and lineage tracing approaches have suggested potential MaSC populations, they have also revealed a number of conflicting results and conclusions.

Multiplicity of mammary stem cells

Development of the mammary gland requires proliferative stem cells that are responsible for the genesis and expansion of the mammary epithelium that drive ductal elongation in puberty and lobuloalveogenesis in pregnancy (known as professional, functional, or *bona fide* stem cells). However, there may also exist a population of cells in the adult mammary gland with the capacity to behave as stem cells under particular conditions (termed facultative or potential stem cells) (Potten and Loeffler, 1990; Visvader and Clevers, 2016). Populations akin to this have been shown in other epithelial systems, but they are less well understood in the mammary gland. These may include either a subset of cells that remain quiescent during normal tissue development, and cells that are recruited under regenerative conditions (Shackleton et al., 2006; Van Keymeulen et al., 2011; Aloia et al., 2016) or in cancer (Koren et al., 2015; Van Keymeulen et al., 2015).

Support for this hierarchal arrangement in the breast that departs from a unidirectional, top-down model is given by transplantation studies. As discussed above, although it is now generally accepted that this method is not suitable for identification of homeostatic stem cells, the underlying experimental observation - that only a certain population of cells is capable of regeneration of the ductal tree - points to the existence of a population of cells that have an intermediate or plastic nature. However, the fact remains that the physiological and pathological role of these cells and their relationship to putative populations of quiescent MaSCs is not immediately apparent.

A pool of quiescent stem cells, which have temporarily and reversibly exited the cell cycle, has been observed in several self-renewing epithelial tissues, including the skin (Cotsarelis et al., 1990) and intestine (Buczacki et al., 2013). These cells may be able to re-enter the cell cycle when required, for example upon injury (Ito et al., 2005) or homeostasis (Boras-Granic et al., 2014) to regenerate the tissue. Of note, these quiescent stem cells are unlikely to be detected by conventional lineage tracing approaches, which require proliferation in order to be able to identify clones (Li and Clevers, 2010).

As such, label-retention assays have been developed for the analysis of slow-cycling and quiescent cells, based on the idea that proliferating cells will dilute out the label. Consequently, any cells that remain labelled after a predetermined chase, known as label-retaining cells, are presumed to be slow-cycling/quiescent stem cells - although they may also be long-lived terminally-differentiated cells. Alternatively, label retention will also occur if the proposed stem cell has retained its template strand of DNA and undergone asymmetric cell division – a theory based on the Cairns/Immortal Strand Hypothesis (Potten et al., 1978). This proposed method of DNA replication allows for long-lived stem cells to retain the mother strand and thus protect against DNA replication-related mutagenesis, with any mutations occurring in the daughter strand passed on to terminally differentiated daughter cells (Cairns, 1975).

A number of different methods have been developed to investigate label retention. Firstly, DNA nucleoside analogues (such as bromodeoxyuridine (BrdU), 5-ethynyl-2′-deoxyuridine (EdU), and [^{3}H]-thymidine) become incorporated into DNA during the synthesis (S) phase of the cell cycle, and as such can be used to label cells that are in cycle at the time of the pulse (Salic and Mitchison, 2008). Indeed, this method has been applied in the mammary gland, which revealed evidence of asymmetric cell division in putative MaSCs using sequential administration of [^{3}H]-thymidine and BrdU labelling (Smith, 2005). This method of cell division would be beneficial for putative MaSCs due to the potential number of cell cycles required to maintain long term regeneration potential with pregnancy and oestrous cycling, without increasing mutagenic load.

Alternatively, genetically modified models can be used to label specific populations of cells, without the requirement for cells being in cycle at the time of labelling. For example, the GFP-labelled histone H2B model allows the expression of H2B–GFP to be temporally moderated by administration of doxycycline (Tumbar et al., 2004; dos Santos et al., 2013). Indeed, the

application of the H2B–GFP model to the mammary gland identified a novel population of cells with an enhanced repopulating ability upon transplantation, marked by the cell surface receptor CD1d (dos Santos et al., 2013). Interestingly, $CD1d^{+}$ mammary repopulating cells were also enriched for Bcl11b expression, a C2H2 zinc finger transcription factor that has recently been independently shown to be associated with physiological quiescence and superior repopulating activity under transplantation conditions (Cai et al., 2017).

However, neither *Cd1d* nor *Bcl11b* mRNAs were found to be enriched in a recently identified quiescent population of basal cells defined by Lgr5 and Tspan8 expression (Fu et al., 2017). These $Lgr5^{+}Tspan8^{hi}$ basal cells, located within the proximal ductal tree, were also demonstrated to have enhanced repopulating activity in limiting dilution transplantation assays. Taken together, these data suggest there may be significant multiplicity within the stem and progenitor population of the mammary gland, even within the putative subset of quiescent mammary repopulating cells.

In light of the ongoing debate regarding the identity and potency of MaSCs, it is reasonable to suggest that there may be a residual population of quiescent bi/multipotent MaSCs that remain in the postnatal mammary gland after embryonic development (Boras-Granic et al., 2014). Given the fundamental necessity for cell proliferation for clone detection in lineage tracing studies, combined with the idea that quiescent stem cells may reside at the apex of tissue hierarchies, this cannot be ruled out (dos Santos et al., 2013).

Indeed, *in utero* DNA labelling has provided some support for this hypothesis, identifying long-lived label-retaining cells that are able to reversibly re-enter the cell cycle and contribute to tissue development and maintenance (Boras-Granic et al., 2014). However, more recent saturation lineage tracing, which labels at least 95% of all cells within a single lineage, has indicated that quiescent MaSCs are lineage-restricted (if they exist and participate in any way to tissue development or homeostasis) (Wuidart et al., 2016). Moreover, analysis of cell division kinetics and telomere lengths in different mammary epithelial populations also suggests that that each lineage is maintained by its own precursors throughout reproductive life (Giraddi et al., 2015).

Given the multiple cycles of regeneration in the mammary gland, as outlined above, the proliferative requirement of mammary stem and progenitor cells throughout reproductive life

is substantial. As such, the relative importance of these putative quiescent MaSCs in normal development and homeostasis remains unclear. However, several important aspects are still to be deciphered: the mechanism by which these quiescent and potential stem cells may be recruited, and by which specific signals in the microenvironment, and their relationship, if any, to the functional stem cells responsible for ductal and alveolar morphogenesis. However, given the complex cellular heterogeneity in breast cancer (Cassidy and Bruna, 2017), a long-lived and highly plastic stem cell could serve as a potential cell of origin for this disease (Blanpain, 2013; Zomer et al., 2013). This highlights the importance of determining the full landscape of MaSC populations and the factors regulating their recruitment.

A potential mammary stem cell niche

It may be the case that potential or quiescent MaSC are not necessarily a separate population of cells but in fact a function of their environment – for example cells may be able to re-enter the cell cycle in response to niche signals, such as hormonal cues. The ability of MaSCs to rapidly and faithfully respond to developmental and homeostatic demands throughout reproductive life may be attributable to their intimate association with a specific cellular microenvironment, known as the mammary stem cell niche. Stem cell niches have been shown to embody discrete and highly specialized sites in particular tissues, for example the crypt base of the small intestine and the hair follicle (O'Brien and Bilder, 2013).

However, other tissues, including the post-pubescent mammary gland, prostate, and lung, seem to lack an easily discernible niche. It has therefore been suggested that stem cells in these organs may instead respond to more ubiquitous tissue signals, and the niche may represent a much more dynamic, less physically defined area. Nevertheless, reciprocal interactions between putative MaSCs and their mature epithelial progeny, neighbouring stromal cells, and the supporting extracellular matrix would undoubtedly provide a combination of autocrine, juxtracrine, and paracrine signals that will direct and adjust cell fate (Sreekumar et al., 2015). Extrinsic regulatory cues may include diffusible molecules (e.g., growth factors and cytokines) as well as mechanical forces (e.g., cell–cell and cell–matrix interactions) (Howard and Lu, 2014).

The absence of conclusive markers, or a molecular portrait of MaSCs, combined with uncertainties regarding their exact location within the post-pubescent mammary epithelium, has greatly impeded the analysis of potential MaSC niches. Cell-surface signatures utilised in

FACS, which have facilitated the isolation of mammary repopulating cells, have provided little insight into the tissue-positional cues that may direct cell behaviour due to the dissociation of the tissue required for cell sorting. However, early transplantation and ultrastructural studies did imply that these mammary repopulating cells were distributed throughout the ductal epithelium (Hoshino, 1962; Daniel et al., 1968; Young et al., 1971), positing that MaSC niches may reside in a 'suprabasal' location in the epithelial bilayer (Smith and Medina, 1988; Chepko and Smith, 1997; Chepko and Dickson, 2003).

Within the pubertal mammary gland, it is generally accepted that the TEBs of elongating ducts likely serve as a transient niche during puberty (Sreekumar et al., 2015; Paine et al., 2016). Therefore, examining the signalling events involved in regulating TEB-resident stem cells is a significant step that may yield important insights into the pathways directing MaSC activity and fate, which may also provide relevant suggestions for the nature of the niche in a post-pubescent gland. Within the TEB, cap and body cells are generally considered to be the precursors of mature basal and luminal epithelial lineages, respectively. Furthermore, cap cells have been long hypothesised to represent an enriched population of bi/multipotent MaSCs (Williams and Daniel, 1983; Smith and Medina, 1988; Srinivasan et al., 2003).

Supporting this, recent work has demonstrated that the alternative promoter of the stem cell-associated phosphatase gene *Ship1/INPP5D* (s-Ship) is exclusively expressed in cap cells during puberty. Furthermore, its expression has been shown to correlate with enhanced mammary repopulating capacity in limiting dilution transplantation assays (Bai and Rohrschneider, 2010). In addition, s-Ship-expressing cap cells are strongly associated with the expression of Par3L, a protein related to the cell polarity regulator Par3 which has been shown to be required for MaSC maintenance and ductal morphogenesis during puberty (Huo and Macara, 2014).

However, the bipotency of cap cells has been brought into question by recent work using mathematical modelling of mammary ductal elongation, suggesting that inwardly migrating cap cells in fact do not contribute to the luminal epithelial lineage as previously hypothesised (Paine et al., 2016). Consequently, the precise contribution of these anatomically distinct cells to ductal morphogenesis requires further investigation. Moreover, the relationship between cap cells in the TEB and the unipotent MaSCs that have been identified by genetic lineage-tracing is also unclear. An answer to these important questions, and a potential unifying definition of

physiological MaSC potency, may be obtained by future inducible fate-mapping studies using transgenic s-Ship and/or Par3L reporter models.

In the post-pubescent mammary gland, where TEBs have fully regressed, the location of MaSCs and their niche constituents is even more ambiguous. One hypothesis is that MaSCs are left behind by elongating TEBs during pubertal growth and consequently are dispersed throughout the adult epithelial network. Here, hormonal cues can then stimulate further branching and the formation of alveolar-like buds and lobuloalveoli during oestrous cycling and in pregnancy, respectively (Brisken and O'Malley, 2010). The notable absence of hormone receptors in mammary repopulating cells (Sleeman et al., 2007) and in MaSC-enriched basal cell populations (Asselin-Labat et al., 2006) implies that paracrine interactions between hormone receptor-expressing cells and stem cells guide tissue development and homeostasis (Asselin-Labat et al., 2010; Joshi et al., 2010; Cai et al., 2014; Rajaram et al., 2015).

Indeed, multiple paracrine signalling pathways are reported to regulate MaSC function downstream of hormone action, including Wnt, EGFR, IGFR, and RANK signalling (Cowin and Wysolmerski, 2010). In addition, FGF, Hedgehog and Notch signalling pathways have also been implicated in modulating MaSC fate during different stages of mammary gland development. However, how the local activities of these pathways are controlled by systemic changes in hormone levels remains unknown (Brisken and Ataca, 2015). Nonetheless, the widespread distribution of hormone receptor-positive cells throughout the adult mammary epithelial tree suggests that MaSCs would be able to receive and integrate these paracrine signals at most architectural locations within the ductal epithelium. Moreover, alterations in the abundance and distribution of hormone receptor-positive cells with age (Ismail et al., 2002) may reflect lifetime-dependent variations in a putative MaSC niche.

MaSCs are thought to survive tissue remodelling during post-lactational involution, enabling further cycles of expansion with each subsequent pregnancy. As such, it may be that MaSCs could reside in the vicinity of epithelial branch points, poised to generate the lateral branches and lobuloalveolar structures required for lactation. Fate-mapping studies using an alveolar-specific whey acidic protein (WAP)-driven Cre have also identified a population of long-lived parity-identified mammary epithelial cells (PI-MECs) that are sustained through multiple reproductive cycles (Wagner et al., 2002). These cells reside at the ductal extremities in the post-parous mammary gland and contribute exclusively to the hormone receptor-negative

luminal lineage in subsequent pregnancies (Chang et al., 2014). Thus, these observations may also support a model wherein an alveolar stem cell niche is positioned near bifurcation sites in the mature ductal epithelium. Interestingly, increased MaSC activity during pregnancy correlates with the re-expression of s-Ship specifically in basal cells at the tips of alveolar buds, suggesting the emergence of a transient stem cell niche during lobuloalveologenesis (Bai and Rohrschneider, 2010).

As outlined earlier in this chapter, distinct adult MaSCs are postulated to fulfil the proliferative and homeostatic demands of the mammary gland. The degree to which this heterogeneity in the MaSC compartment is intrinsic or instead a result of microenvironmental cues, however, is not known. A recent single-cell lineage-tracing study, which employed quantitative volumetric analysis to determine the contribution of a single labelled MaSC to ductal morphogenesis, estimated that at least 35 lineage-restricted MaSCs actively and stochastically contribute to the development of each major duct during puberty (Wuidart et al., 2016). A subsequent study, also using quantitative lineage-tracing at clonal density, put this number at 260 lineage-restricted MaSCs per TEB, leading to the suggestion that most TEB cells can function as lineage-committed MaSCs (Scheele et al., 2017).

Discrepancies between these two studies may reflect differing functional definitions of MaSCs as well as the quantitative and mathematical platforms and assumptions for analysis. Quantitative lineage-tracing studies also suggest that molecularly heterogeneous populations of TEB-resident MaSCs function as single equipotent pools, colonising ductal branches through stochastic neutral drift dynamics (Scheele et al., 2017). Random segregation during successive rounds of TEB bifurcation mediates the unequal distribution of MaSC progeny between adjacent ductal structures, leading to clonal enrichment or extinction over time, supporting previous observations of clonal labelling patterns. These early applications of quantitative and single-cell lineage-tracing approaches in the mammary gland have provided unprecedented insights into clonal dynamics and stem/progenitor heterogeneity and multiplicity, heralding a new era in our investigation and understanding of normal and malignant stem cells in the breast.

A novel approach for lineage tracing in the mammary gland

Previous lineage tracing studies in the mammary gland have facilitated *in situ* examination of MaSC properties under conditions of minimal interference. However, unlike transplantation

assays, a limitation of these studies is that they have traditionally been unable to map the fate of a single labelled cell (Shackleton et al., 2006). As such, due to the inherent biases of these two approaches discussed above, a neutral way to assess stem cell dynamics in a physiological context, that makes fewer prior assumptions and can be interpreted at various stages of development, is required to resolve the long running debate on the MaSC hierarchy.

Consequently, for a novel lineage tracing approach, independent of pre-defined, promoter driven reporters, the work presented in this thesis utilises two transgenic mouse strains to neutrally trace single stem/progenitor cells. The first, the $R26^{[CA]30}$, exploits the inherent instability of microsatellite repeats to cause spontaneous, random frame shift mutations during DNA replication. For this, a $[CA]_{30}$ dinucleotide repeat tract is placed upstream of an out-of-frame reporter cassette (either enhanced yellow fluorescent protein (YFP) or a modified, thermophillic beta-glucosidase (SYNbglA) (McCutcheon et al., 2010; Kozar et al., 2013). During DNA replication, a mismatch repair (or "slippage") of the reading frame can randomly occur, potentially moving the reporter in-frame and thus triggering constitutive reporter expression in the "slipped cell" from the constitutively expressed Rosa26 locus. Accordingly, this leaves an indelible mark on the cell and, if this is a stem/progenitor cell, all of its progeny will also be marked. Importantly, this slippage event is exceedingly rare, and thus it can be used to track the fate of a single labelled cell with confidence

Because terminally differentiated cells are unable to divide, subsequent analysis of the clone size, structural organisation and cellular composition of clonally-marked regions in the mammary gland can provide useful insights regarding the nature of mammary stem/progenitor cells - including potency, location and capacity. Moreover, because strand slippage is more likely to occur in proliferating cells, this model will be used to look at the active contribution of the proliferative, functional stem cells that primarily contribute to the development of the gland during puberty, and during the highly proliferative pregnant phase resulting in a lactating mammary gland.

However, for the ability to time tracing and to enable marking of quiescent cells another model must be used. For this, R26R-Confetti mice will be employed (Livet et al., 2007; Snippert et al., 2010). Using a similarly neutral approach to the $R26^{[CA]30}$ model, mice that express inducible Cre recombinase in all cells ($R26^{CreERT2}$) are crossed with $R26^{Confetti}$ animals to achieve low-density, multicolour and unbiased labelling.

In this latter model, the constructs contain two tandem invertible DNA segments; with administration of tamoxifen, Cre recombinase is translocated to the nucleus and removes the loxP flanked "STOP" roadblocks. Subsequent stochastic inversion and excision recombination events create four random expression possibilities (nuclear green fluorescent protein (GFP), cytoplasmic YFP, cytoplasmic red fluorescent protein (RFP), or membrane-bound cyan fluorescent protein (CFP)). Cells will thus be marked at random in one of four colours. Importantly, any cell can be labelled with this model because cells are traced under the ubiquitous Rosa26 promoter. Furthermore, labelling density can be titrated, with sparse reporter induction attained using low doses of tamoxifen. It also provides the benefit of a multicolour reporter, which aids with the distinction of individual clones.

1.3. Imaging the murine mammary gland

The location of the stem/progenitor cells in the mammary gland is another important aspect that requires clarification. Previously, proposed identities of MaSCs have been derived from FACS sorted cells, which removes important spatial information, or lineage-traced cells, which are visualised from 2D sections and may not accurately reflect the true nature of the heterogeneous gland, and fail to provide information about spatial distribution of clones (Sale and Pavelic, 2015). Moreover, this becomes especially relevant in the context of single cell genetic lineage tracing, whereby labelling is rare and as such requires analysis of the whole tissue which would not be readily achievable using thin 2D tissue sections.

Historical approaches

Given this, 3D imaging in the mammary gland has been previously demonstrated by utilising a number of different methods. Indeed, fluorescent imaging using confocal microscopy has shown success in imaging the mammary primordia during embryonic development (Kogata and Howard, 2013). However, the inherent smaller sample size associated with embryos mean they do not face the same challenges as the postnatal mammary gland, such as deeper imaging depths and a higher proportion of adipose tissue, and as such this technique is not compatible with the study of the gland after embryonic development.

Within the pubertal gland, protocols for colourimetric wholemount imaging - such as X-gal labelling for LacZ - are often not compatible with high resolution imagining. This means that

imaging is usually carried out using stereomicroscopy which renders the visualisation of individual single cells not possible (Sale and Pavelic, 2015). To combat this, an improved X-gal labelling protocol has been developed for better imaging of single cells within wholemount mammary glands, which allows for high magnification imaging at different focal depths (Šale et al., 2013). Using this method combined with a Notch2 paralogue as a genetic marker, the authors were able to identify two previously unknown subpopulations of luminal cells with novel topological locations which would not have been detected in 2D using traditional serial sections. However, this method can still not be combined with *in situ* co-staining, meaning it had to be supplemented with an additional fluorescent model which allowed for definitive identification of cell types.

Consequently, another method that utilises fluorescent multiphoton imaging has been used, that even demonstrated live imaging of mammary glands by taking advantage of the glands inherent autofluorescence (Johnson and Mueller, 2013). Furthermore, use of multiphoton imaging provided the additional benefit of using second harmonic generation (SHG). This process allows for imaging of collagen fibres based on their structural signal, without the need for additional staining. Interestingly, the authors also combined this with imaging wholemount glands from HAI-1 transgenic mice, which overexpresses the Kunitz-type protease inhibitor HAI-1 in their mammary glands, that had been previously stained with Carmine Alum. The wholemount glands were able to be retrospectively imaged, with SHG used to image collagen fibres and the fluorescent signal from the Carmine Alum stain utilised to detect the epithelial ductal tree. This identified abnormal TEB and collagen fibre development which could not be fully appreciated in H&E stained paraffin sections, thus highlighting the utility of analysing tissues in multiple focal planes in the context of disease. However, this protocol faces similar issues in that the light refractive adipose tissue impeded successful high-resolution imaging, and only structures relatively close to the surface of the tissue were able to be imaged. Moreover, this protocol was still restricted in the options for staining, limiting its utility.

A number of more recent protocols have therefore been developed to allow labelling of multiple aspects, given the necessity to be able to differentiate between at least the luminal and basal epithelial compartments. The first involves the use of microdissection, whereby portions of the epithelial ductal tree are removed from its native stroma (Rios et al., 2014). This therefore negates the need to image through the large volumes of adipocytes and other stromal cells that are present in the gland and are especially light refractive and as such limit

the imaging depth, before even reaching the epithelial compartment of the gland (Rios et al., 2014). The use of this technique was combined with lineage tracing which allowed imaging of large clonal regions in the ductal network (Rios et al., 2014; Wang et al., 2015; Fu et al., 2017). However, this method comes with a number of drawbacks. Firstly, it limits the extent of epithelium that can be imaged, with often only small sections of the ductal network being able to be dissected. Along with this, there is a requirement for a high degree of technical skill to carry out dissection on such a small area. Moreover, it rules out the use of 3D imaging to study the native stromal compartment of the mammary gland. Together, this means that the use of microdissection for 3D imaging in the mammary gland may miss out important aspects.

Another method previously used for 3D imaging of the mammary gland involves a combination of enzymes, such as collagenase and hyaluronidase, to dissolve the extracellular matrix – and in the case of the mammary gland – the surrounding adipocytes (Wuidart et al., 2016). Consequently, the mammary gland can then be imaged *in toto.* However, achieving specific digestion of the stroma whilst also leaving the epithelium intact is a delicate balance. Consequently, recent work has shown that the use of this proteolytic digestion in the mammary gland has a detrimental effect on the structure of the epithelium (Rios et al., 2016). Specifically, the basal cell layer was shown to be particularly affected, presumed to be due to the destruction of the basal lamina on which basal cells rely on for cell contact. Using even low concentrations of proteolytic enzymes resulted in depletion of cells in the outermost layer of the ductal tree, leading to gaps in the usually continuous morphology. Furthermore, the basal cells which remained, displayed an altered morphology, losing their traditional elongated shape and becoming more rounded (Rios et al., 2016). This may have important implications when using this technique to assess the results from fate mapping studies – indeed, it has been postulated that rare, bi-lineage clones may be missed under these conditions if certain cell types are preferential affected by the protocol (Rios et al., 2016)

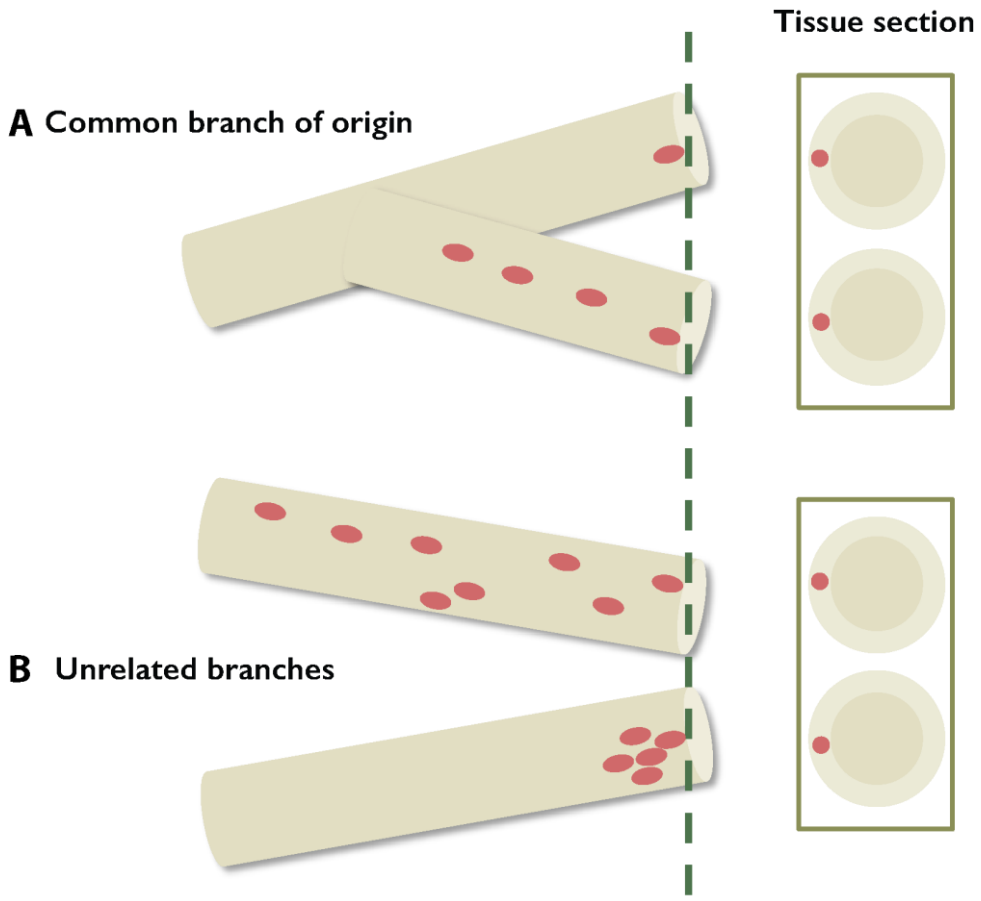


Figure 1.3 Limitations of 2D imaging in the mammary gland

The complex topological architecture associated with a high branched structure such as the mammary gland presents issues when imaging in two focal planes. Schematic shows representative labelled cells from lineage tracing studies and their associated 2D tissue sections. **(A)** Shows cells arising from presumably the same mother cell, with a common branch of origin. **(B)** Shows labelling from two unrelated branches and thus potentially unrelated progeny. However, both 2D sections would reveal the same result, thus confounding interpretation of lineage tracing data, especially in the case of rare or clonal labelling. The 2D sections would also not reveal the differential labelling patterns. Figure adapted from (Sale and Pavelic, 2015).

A new technique for imaging the mammary gland

Altogether, this highlights the importance of imagining the mammary gland in multiple focal planes, and as such the development of methods to achieve this that are without the limitations of those previously used. Consequently, this thesis will describe the development and optimisation of optical tissue clearing methods for use in the mammary gland. This process involves treatments to strip away lipids and/or match the refractive index of tissues and thus minimise the scattering of light at the oil-water interface. This is particularly important in the virgin mammary gland, which is difficult to image at depth due to the high proportion of fat, which is especially light-refractive (Susaki et al., 2014). As a result, optical clearing permits the visualisation of mammary tissue to a single cell level, while allowing for staining and retaining the fluorescence (or colourimetric) signal, tissue integrity and important spatial information.

Optical clearing involves use of strong detergents and matching of refractive indices in order to enable imaging of whole mount organs and tissues. First established for use in the brain, a number of different methods have now been developed that have allowed novel insights into the 3D architecture of tissues. For the mammary gland, the ability to image the tissue in its entirety is particularly important due to its unpredictable ductal architecture, which presents unique challenges in terms of interpretation of lineage tracing data. Traditional tissue sections can be insufficient for accurate quantitative analysis and do not reveal complex spatial patterns (summarised in Figure 1.3). As such, methods to provide the ability to image the gland in 3D are vital. Indeed, traditional use of the wholemount stain, carmine alum, has been utilised extensively in mammary gland research despite its incompatibility with other antibody co-staining - demonstrating the importance for researchers to be able to visualise the entire ductal tree in the mammary gland.

As such, optical tissue clearing provides a method to yield the benefits of imaging the entire mammary gland in 3D, without detrimental effects on either the stroma or the epithelium. Imaging the ductal network, *in situ* and without alteration, removes potential complications or anomalies that arise and that may contribute to disparities in the interpretation of data in the field of mammary gland research.

1.4. Aims of this thesis

Despite research in mammary gland biology spanning over 60 years, definitive elucidation of a number of the aforementioned key aspects of mammary gland biology remain. The techniques and associated limitations previously used have unfortunately hampered conclusive definitions in these important areas. Consequently, the work described in this thesis aims to develop and optimise methods for use in the murine mammary gland to allow study of its tissue development and stem cells. This will be done while avoiding the shortcomings associated with the techniques which have been traditionally applied to mammary gland research. As such, the primary objective of this work is to further characterise the requirements for MaSCs during normal tissue morphogenesis, and maintenance in the mature gland. To achieve this, this thesis has the following overall aims:

Aim 1 : Develop 3D imaging methods for the mammary gland

The utility of analysing the mammary gland in multiple focal planes has been well demonstrated. Consequently, the use of optical tissue clearing in the mammary gland will be assessed by testing multiple protocols and optimising them accordingly. Moreover, the imaging techniques to visualise the tissue in 3D will be developed. These methods will then be applied to the work in the following chapters.

Aim 2 : Lineage tracing in the mammary to further understand MaSC characteristics

The use of neutral lineage tracing models to study cell fate in the multiple stages of mammary gland development will be established and assessed. These will be used to identify clonally-marked regions in the embryonic, pubertal, adult and lactating mammary gland to elucidate the nature of MaSCs, in terms of their potential, capacity and location. Based upon previous work, the hypothesis is that multiple unipotent adult MaSCs contribute to ductal morphogenesis in pubertal development and alveolar development in pregnancy. It is further hypothesised that these are present throughout the postnatal mammary gland in architecturally-restricted sites. This will be examined by analysing the mammary gland in 3D for full appreciation of the clonal characteristics.

Aim 3 : Identify the proliferative capacity of the mammary gland

Next, the proliferative capacity of the mammary gland during different stages will be assessed. This will be done in pre-pubertal, pubertal, young adult and finally older adult mammary glands. Combining this knowledge with the results from Aim 2 will provide insights into the functional requirements and potential locations of MaSCs.

Aim 4 : Assess the use of an *ex vivo* system of mammary gland development

Finally, the work carried out in the previous three chapters will be combined to assess the use of *ex vivo* 3D organoid models of the mammary gland to study development. Specifically, fate mapping of the cells will be carried out using cells from the lineage tracing models, and proliferation assessed and compared with these results. Furthermore, use of optical clearing will be tested for imaging the organoids in three dimensions.

1.5. Importance and implications

The mammary gland provides a unique opportunity to investigate epithelial development extra-embryonically that is not available in other tissues. Moreover, study of maintenance and turnover of this organ has important implications for other epithelial systems. As such, combining methods for non-proteolytic 3D imaging, together with quantitative platforms for image analysis which consider tissue architecture, cell morphology, chimerism, and Cre-specificity will aid the study of mammary gland development and the contribution of stem cells, in addition to potentially providing new insights for epithelial tissue development.

Elucidation of the normal MaSC hierarchy also has important implications for understanding the biology of the mammary gland in relation to neoplastic development. The heterogeneity of breast cancer is complex and although there is now an increased understanding of molecular and pathological subtypes, there is still much work required to get closer to understanding the cell(s) of origin of breast cancer. Given the proposed longevity and suggested ability of MaSCs to survive multiple waves of cell death in involution, they represent a logical candidate for a potential cell of origin of breast cancer. Reponses to treatment are still highly unpredictable and the hope is that better stratification of breast cancer can improve response to therapy and may lead to the development of more targeted therapies. By increasing our understanding of the potential cell(s) of origin, this work could enhance current methods of patient stratification and understanding of breast cancer development.

Chapter 2

Materials and methods

2.1 Mice

2.1.1 Husbandry

Experiments using animals were conducted under a United Kingdom Home Office license in accordance with the *United Kingdom Animals (Scientific Procedures) Act 1986* and in accordance with *European Union Directive EU 86/609*. Specific experimental studies were approved by the University of Cambridge Department of Pathology animal unit and animals were housed in individually ventilated cages under a 12:12 hour light–dark cycle, with water and food available *ad libitum*.

All non-pregnant/lactating animals were assessed for oestrus cycle staging based on a non-invasive protocol in order to prevent unnecessary stress and complications with pseudo-pregnancies as previously described (McLean et al., 2012). Briefly, after visual assessment, approximately 100 μL of sterile phosphate buffered saline (PBS) was used to carry out a vaginal lavage, which was then smeared onto a glass slide and left to air dry. This was stained with toluidine blue and the oestrus stage was assessed based on cytological assessment as visualised under a light microscope, with the stage categorised depending on the presence and/or proportion of nucleated epithelial cells, cornified epithelial cells and lymphocytes. For studies during lactation, mice were mated with studs, allowed to litter and tissue was harvested between lactation days 2 to 10.

For tissues, female mice were sacrificed by dislocation of the neck or terminal anaesthesia. Eight mammary glands (excluding the first (cervical) pair) were dissected and immediately spread and fixed on card (Tetra Pak) in 10% neutral buffered formalin (NBF) for 9 hours at room temperature for all experiments, with the exception of EdU incorporation studies where tissues were fixed overnight at 4°C. All animals were assessed also for oestrous cycle staging posthumously, as previously described. Fixed tissue was then stored at 4°C in PBS containing sodium azide (0.05% (w/v)) for up to 8 weeks.

2.1.2 Mouse models

The R26$^{[CA]30}$SYNbglA and R26$^{[CA]30}$YFP strains were a gift from Doug Winton and have been previously described (Kozar et al., 2013). Briefly, [CA30]-SYNBglA and [CA30]-YFP reporter cassettes were synthesized by Yorkshire Bioscience. These were subsequently subcloned by

Nhe1 digestion into the pROSA-MCS13-puro targeting construct, carried out by the transgenic facility at the Cancer Research UK London Research Institute in an embryonic stem cell line of C57BL/6J genetic background (cell line: B6 1.1). $R26^{[CA]30}$ SYNbglA and $R26^{[CA]30}$ YFP mice were then maintained as hemizygotes on a C57BL/6J genetic background. Animals were examined at 7 weeks of age for all studies in pubertal mice. All quantitative analysis was performed on mice that were hemizygous for $R26^{[CA]30}$SYNbglA or $R26^{[CA]30}$YFP. For genotyping primers and conditions see appendix. R26R-Confetti mice have been previously described (Livet et al., 2007; Ventura et al., 2007) and were a gift from Jenny Nichols. Mice that were hemizygous for both R26-Confetti and R26-$CreER^{T2}$ were generated by mating homozygous mice. For genotyping primers and conditions see appendix.

2.1.3 Initiation of lineage tracing

Cre-mediated recombination was achieved through a single intraperitoneal injection of 0.5 mg tamoxifen in virgin mice, or oral gavage of 33.3 mg/kg in pregnant mice unless otherwise stated. Lineage tracing was induced at 4 weeks for studies in pubertal R26-Confetti;R26-CreERT2 mice, unless otherwise stated. Tissues were collected after a 2-day chase to determine initial labelling or after 3 weeks for pubertal lineage-tracing studies. Tamoxifen was made up in sterile corn oil.

2.1.4 Tumours

Syngeneic mammary tumours were established by orthotopically implanting 5×10^3 TUBO cells, into the abdominal (fourth) mammary gland. This cloned cell line was established from a mammary carcinoma that spontaneously arose in a BALB-neuT mouse and therefore carries the Her-2/neu oncogene driven by the MMTV promoter, as previously described (Rovero et al., 2000). Mice were monitored regularly, and tumours were harvested before exceeding humane endpoints (approx. 4–5 weeks from implantation).

2.1.5 EdU incorporation studies

EdU powder (life technologies, catalogue number C10339) was reconstituted in sterile saline at a final concentration of 10 mM, according to manufacturer's guidelines. Female mice (aged 4, 7, 12 and 24 weeks, n=10 for each) were oestrous staged at the same time each morning for one week prior to injection (described above), and when in dioestrous were administered 25mg/kg via I.P. injection. A non-injected saline-control mouse was included in each cohort. 4 hours after injection the mice were sacrificed by a Schedule 1 approved methods (typically

cervical dislocation), and upper and lower mammary glands were dissected, along with the first 5 cm of the small intestine as proliferative positive controls, in addition to vaginal lavage samples taken for oestrus staging confirmation.

2.2 Histology

2.2.1 SeeDB optical clearing

See deep brain (SeeDB) optical clearing was adapted for the mammary gland from the protocol outlined by Fujimoto *et al* (Ke et al., 2013). Samples were permeabilised and blocked in 10% BSA with 1% Triton-X in PBS, overnight at 4°C on a shaker. The following day primary antibodies were added, diluted in blocking solution, and samples incubated for 4 days at 4°C on a shaker. Samples were then washed 3 times for 1 hour on a shaking platform at room temperature with PBS, before appropriate secondary antibodies added and samples again incubated at 4°C on a shaker for 2 days. They were then washed again in PBS at room temperature and stained with 10uM 4',6-diamidino-2-phenylindole (DAPI) diluted in PBS for up 3 hours at room temperature. Following this, the samples were put through increasing concentrations of fructose with 0.5% α-thioglycerol every 12 hours (20, 40, 60 then 80%), at room temperature on a rocking platform. Finally, the samples were left in 100% fructose for 24 hours followed by SeeDB solution (115% w/v) for another 24 hours and then finally imaged using confocal microscopy.

2.2.2 CUBIC optical clearing

CUBIC clearing in the mammary gland was adapted from method outlined by (Susaki et al., 2014). Tissues were immersed in reagent 1a for 24 hours at 37°C (a less harsh adaptation of the original reagent 1, unpublished with protocol available at http://cubic.riken.jp/), urea (10% w/w), N,N,N',N'-tetrakis (2-hydroxypropyl) ethylenediamine (5% w/w) and Triton X-100 (10% w/w) with 5M NaCl in distilled water). The next day samples were permeabilised and blocked with incubation in 10% normal goat serum (NGS) with 0.5% Triton-X in PBS, overnight at 4°C on a shaker. The following day appropriate primary antibodies were added and samples incubated for a further 4 days at 4°C on a shaker. Following this, samples were washed 3 times for 1 hour on a shaking platform with PBS + 0.1% Triton-X at room temperature, before appropriate secondary antibodies added and samples again incubated at 4°C for 2 days on a shaker. Samples were then again washed and incubated for up to 60

minutes with 10uM DAPI then finally left in reagent 2 (sucrose (44% w/w), urea (22% w/w), 2,2',2''-nitrilotriethanol (9% w/w), Triton X-100 (0.1% w/w) in distilled water) for 24 hours at 37°C on a shaker before being imaged. For full CUBIC regent recipes see appendix.

2.2.3 PACT-sRIMS optical clearing

The A4P0 hydrogel formulation was selected for PACT- based clearing of the mouse mammary gland. A4P0 was prepared to contain acrylamide (4% (v/v)), 2,2′-Azo- bis[2-(2-imidazolin-2-yl)propane] dihydrochloride (0.25% (w/v)) in PBS. PACT-clearing solution consisted of SDS (8% (w/v)) in distilled water, pH 7.5. Mammary tissue pieces (approx. 10 × 10 × 1 mm) were incubated in A4P0 hydrogel monomer for 4 days at 4°C and heated to 37°C in a water bath for 4–6 h. Excess gel was carefully re- moved from the tissue and samples were immersed in PACT clearing solution for 24 hours at room temperature. Samples were immersed in fresh clearing solution, incubated at 37°C for 4 days (with replenishment every second day), and finally washed with PBS containing triton- X100 (0.1% (w/v)) for 24 h. For immunostaining, samples were blocked in PBS containing triton-X100 (0.5% (w/v)) with goat serum (10% (v/v)) overnight at 4 °C. Primary antibodies were diluted in blocking buffer at 4°C for 4 days with agitation. The tissue was washed 3 times for >1 hour in PBS then incubated with Alexa Fluor conjugated secondary antibodies for 2 days at 4°C with agitation. Samples were then again washed in PBS and incubated with DAPI (10 μM) for 2–3 h. samples were incubated in sRIMS for 4 days or until imaging. sRIMS was prepared by combining sorbitol (70% (w/v)) in 0.02 M phosphate buffer (Marx, 2014; Yang et al., 2014).

2.2.4 Optical clearing and measurement of sample size changes

Mammary tissue pieces were processed using 3DISCO, PACT-RC, PACT-sRIMS, CUBIC or SeeDB-based tissue clearing protocols, and images were acquired on a dissecting microscope (Leica MZ75) with constant exposure, gain and magnification. For quantification of sample size changes, image thresholding was performed using ImageJ (v1.49p, National Institutes of Health) and the pixel area was measured. Volume changes were calculated as the ratio of the pixel area before and after tissue clearing.

2.2.5 Human tumour optical clearing

Optical clearing of human tumour samples and collagen scaffolds was carried out by Dr Robert Hume (Hume et al., 2018a). The human breast biopsy material was provided by the Breast Cancer Now Tissue Bank held at Barts Cancer Institute, Queen Mary University of London,

United Kingdom. This tissue bank is licensed by the Human Tissue Authority, according to United Kingdom ethical guidelines and approval, and biopsy material is obtained following informed consent. Anonymized data linked to each sample are provided.

Human ER+ breast tumour biopsies were fixed overnight in 4% paraformaldehyde (PFA) at 4C followed by immersion in Reagent 1a (modified from Reagent 1, unpublished, available at http://cubic.riken.jp) at 37°C for 3 days changing into fresh Reagent 1a each day. Samples were blocked overnight in 0.5% Triton-X 10% normal goat serum–phosphate buffered saline (NGS-PBS) (blocking solution). Primary antibodies were diluted in blocking solution and agitated on a rocker for 5 days at 4°C. Samples were washed in PBS briefly followed by 3 x 1 hour washes in fresh PBS. Secondary antibodies were diluted in blocking solution and agitated on a rocker for 2 days at 4°C. Samples were washed in PBS briefly followed by a 1 hour wash in fresh PBS. Nuclei were marked with a 2 hour wash in 10 mM DAPI. Samples were washed in PBS briefly followed by a 1 hour wash in fresh PBS and immersed in Reagent 2 at 37°C in a dry incubator for 24 hours before imaging. Two-photon microscopy techniques were carried out on a LaVision BioTec TriM Scope II upright two-photon scanning fluorescence microscope.

2.2.6 Antibodies

The following primary antibodies were used for immunostaining: rabbit anti-K5 (Covance, PRB160P, 1:100), rat anti-K8 (Developmental Studies Hybridoma Bank, TROMA-I, 1:50), rabbit anti-SMA (Abcam, ab5694, 1:300), mouse anti-SMA (Abcam, ab7817, 1;200), rabbit anti-E-cadherin (Cell Signalling, 3195, 1:50), rabbit anti-PR (DAKO, A0098, 1:50) and chicken anti-GFP (Abcam, ab13970, 1:2000), anti-laminin (1:500; Abcam, ab11575), anti-collagen IV (1:500; Abcam, ab6586), anti-Ki67 (1:100; Abcam, ab15580), anti-ERα (1:50; Leica, NCL-L-ER-6F11).

The following Alexa Fluor-conjugated secondary antibodies were purchased from Life Technologies and used 1:500: goat anti-mouse 488 (A11001), goat anti-mouse 647 (A21237), goat anti-rat Cy3 (A10522), goat anti-rat 488 (A11006), goat anti-rabbit 488 (A11008), goat anti-rabbit 647 (A21245), goat anti-chicken 488 (A11039) and goat anti- chicken 568 (A11041).

2.2.7 EdU detection

EdU incorporation was detected using Click-iT EdU Alexa Fluor 594 or 488 Imaging Kit (Life Technologies, catalogue number C10339), using the manufactures protocol but adapted and optimised for whole mount staining. Briefly, samples were fixed overnight in 10% NBF,

followed by block and permeabilisation overnight in 10% BSA in 1% PBS-Triton-X. Following this, samples were washed in 3% BSA then the Click-iT reaction cocktail was made up according to manufactures protocol and samples incubated with the solution for 2 days at 4°C on a rocking platform. After this, samples were washed then usual steps for primary and secondary antibody staining carried out, as detailed above.

2.2.8 Methyl green staining

Due to the colour of the SYNbglA lineage tracing being similar to traditional carmine whole mount staining, an alternative with different colouring was required. After assessment of several histological stains Methyl green was chosen. Tissues were initially cleared using CUBIC, as outlined above. They were then incubated in 0.5% methyl green solution for 1.5 hours at room temperature, before being briefly rinsed 3 times in H_2O. Tissues were then de-stained for 15-30 minutes in 50 mL of 50% ethanol plus 122.5 μL of 32% hydrogen chloride (HCl), with level of destaining checked regularly, before finally being rinsed in PBS.

2.2.9 Wholemount DAB staining

A similar CUBIC protocol as outlined above was utilised for whole mount diaminobenzidine (DAB) staining with some alterations. Samples were left in reagent 1 for 2 days then reagent 2 for 1 day (both at 37°C), then washed in PBS briefly before being incubated overnight in 3% hydrogen peroxide with 20% DMSO in methanol on order to quench endogenous peroxidase activity. The following day the samples were blocked with 10% NGS in 0.5% Triton-X in PBS, overnight at 4°C, followed by incubation with primary antibodies for 4 days at 4°C. Following this samples were washed then incubated with appropriate species- specific conjugated horseradish peroxidase (HRP, all Dako), for 2 days at 4°C. For DAB staining, samples were washed then DAB (Sigma, catalogue number D5905) added for 30-60 minutes. After being briefly washed, samples were stained with 0.5% methyl green for 1.5 hours. After this, samples were briefly rinsed in distilled H_2O then de-stained for up to 30 minutes in ethanol and HCl. Samples were then put back in reagent 2 for 24 hours before being imaged. Alternatively, quenching, blocking and antibody steps can be performed after immersion in reagent 1 with a similar outcome.

2.2.10 Organoid wholemount imaging

For whole-culture imaging of organoids, media was first removed and wells washed with cold PBS three times. Using widened P1000 pipette tips pre-coated in FBS, the Matrigel was gently dissolved in ice cold PBS and organoids transferred to 15mL falcon tubes. The organoids were left to settle to the bottom, supernatant removed and then fixed in 4% PFA at room temperature for 30 minutes. After this, they were washed three times with PBS for 10 minutes before the blocking / permeabilisation step, using PBS + 5% FBS + 2% BSA + 0.2% Triton for 2 hours at room temperature on a shaking platform. Primary antibody staining was then carried out, with antibodies diluted in blocking buffer, overnight at 4 C. The following day, after three washes in PBS, secondary antibody staining was carried out (again diluted in blocking buffer), for 2 hours at room temperature. After a further three washes in PBS, DAPI staining was carried out for 1 hour. Following a final wash, organoids were then imaged.

2.3 Mammary organoids

Mouse mammary organoids were developed based on the protocol by Jarde *et al.* Female mice were culled by cervical dislocation and 4th and 5th mammary glands dissected, with lymph nodes removed. Glands were placed on ice in L-15 + 10% foetal bovine serum + penicillin / streptomycin + 10% glutamax until use. Following this, tissues were dissociated based on previously described protocols. Tissues were dipped in 70% ethanol and transferred to fresh L-15/FPSG for mincing. To do this, tissues were placed in sterile petri dishes and minced using two scalpel blades until a fine, semi-liquid slurry was formed. Minced tissue (up to 6 mice per 35ml) was then incubated with sterile-filtered trypsin (Sigma, T5266) / collagenase A (Roche, 11088793001) mixture in L-15/PSG media at 37°C for 1 hour with agitation.

Following digestion, tissue was spun (all centrifugations carried out at 1500 rpm for 5 minutes at 4°C unless otherwise stated) and the resulting fat layer was removed and transferred to a fresh falcon tube to be centrifuged again. The original pellet was then resuspended in 1mL of L-15/ FPSG and washed with 9mL of media, then combined with any remaining pelleted cells from the fat layer. The combined pellet was then re-suspended in red blood cell lysis buffer and incubated for 5 minutes at room temperature. The sample was then washed and centrifuged, and the resulting pellet was resuspended thoroughly in 1mL DMEM + 10% FBS + PSG with a P1000 pipette and transferred to a T75 tissue culture flask and incubated for 1 hour at 37°C and 5% CO_2 for depletion of fibroblasts.

Following this, the majority of fibroblasts should have attached and with moderate shaking the epithelial cells were removed. After rinsing off any remaining un-attached cells, the cell suspension was centrifuged and washed in order to remove any remaining serum. The resulting pellet was then gently resuspended in the required volume of growth factor reduced Matrigel (BD, 356231) with pre-cooled pipette tips. 50µL drops were then placed in the centre of wells on pre-warmed 24-well plates and placed in an incubator to set. After 10 minutes, 500µL of defined medium was added to each well (see appendix for composition).

Media was changed three times per week, and organoids were passaged approximately every 14 days, or on an as-needed basis. To do this, media was removed from wells and 500µL of ice cold DMEM/F12 added to each well. Using pre-cooled P1000 pipette tips, the Matrigel was broken up and mixed approximately 15 times. Well contents were then transferred to a 15mL falcon and each well was washed with a further 1mL which was combined with the previous contents. The contents were washed with DMEM/F12 and centrifuged, supernatant removed and washed and centrifuged again to remove any remaining Matrigel.

Following this, the supernatant was removed and dissociation was carried out by resuspending the pellet with 500µL of trypsin-EDTA 0.05% (Invitrogen, 25300054) and mixing with a P1000 pipette 5 times (pre-coated with FBS to ensure organoid structures did not stick to tips). The sample was incubated at 37°C for 5 minutes, then an equal volume of trypsin inhibitor was added and mixed thoroughly 20 times with a P1000 pipette. After waiting 2 minutes for the larger, un-trypsinised structures to sink to the bottom, the supernatant was strained through a 70µM cell strainer (BD, 352350). The trypsinisation and straining process was then repeated until all larger organoid structures were broken down and the cell suspension was centrifuged. After removal of supernatant, the pellet was thoroughly resuspended in the required volume of Matrigel and re-plated onto pre-warmed plates as previously described.

2.4 Microscopy

2.4.1 Confocal

Images of wholemount mammary glands were acquired using a Leica TCS SP8 inverted confocal microscope with 10×/0.4 or 20×/0.75 HC PL APO objective lenses. Laser power, line averaging and step increment were adjusted manually to give optimal fluorescence intensity

for structure depth for each fluorophore with minimal photobleaching. For confetti data, all channels (RFP, GFP, YFP) were acquired regardless of labelling (with exclusion of CFP due to use of DAPI as a nuclear marker). Images were acquired with a pixel intensity range of 12 or 16 bits to obtain the maximum possible illumination information for subsequent data analysis and quantification.

2.4.2 Light sheet

Samples were immunostained and cleared according to the CUBIC protocol. After clearing, samples were embedded within an agarose (1% (w/v) in H2O) tube, prepared by aspirating agarose (37–38°C) into a pre-warmed 1-mL syringe in which the syringe neck had been cut off. Mammary tissue strips were quickly placed within the agarose tube using forceps and centred by rolling the syringe between the palms. After setting, the plunger was removed, and the entire syringe was submerged in CUBIC reagent 2 for 24 hours. Samples were imaged in reagent 2 or glycerol in H2O (34% (w/w)).

The light sheet system was a home-built modified version of the OpenSPIM system (Pitrone et al., 2013). The microscope was built and operated in the T-SPIM layout, whereby illumination happens from two sides simultaneously by overlapping two individual sheets to allow a more even illumination and to reduce artefacts, such as striping. We used two Olympus 5×/0.15 air lenses to generate the light sheet. The higher refractive indices, long working distance (20 mm) and the fact that the lenses were on threaded mounts allowed us to adjust the point of focus accordingly. The imaging light path was equipped with a Nikon 16×/0.8 water dipping lens.

We imaged onto an Andor Neo 5.5 (ANDOR) or a Hamamatsu ORCA- Flash4 V2 (Hamamatsu) with 6.5 µm pixels. For excitation a home-built laser combiner was used, bundling 405 nm, 488 nm, 561 nm and 640 (Coherent Cube 405 and 640, Coherent Sapphire 488 and 561) into a single- mode fibre. Channels were acquired sequentially, and emission was filtered by suitable band-pass or long-pass filters (DAPI: 447/60; AF647: 705/72; both AHF Analysentechnik). The sample was mounted in a 4D (xyzθ) stage (Picard Industries) allowing optimal positioning of the sample in the light sheet. During imaging the sample was moved through the light sheet with a step size of 1.5 µm and the light sheet thickness was adjusted to be 6 µm to warrant an even thickness of the sheet across the entire sample width. Exposure times were between 15 and 150 ms.

2.4.3 Stereoscope

For wholemount bright field transmission imaging of mammary glands, images were acquired using a Leica MZ75 dissecting microscope with Leica Acquisition Software (LAS).

2.5 Imaging analysis

2.5.1 Imaris

Images were processed and 3D image reconstructions were generated using Imaris image management software (v8.0, Bitplane) or Image J Fiji (version 2.0, National Institutes of Health) using 3D and / or maximum intensity projections. Denoising of 3D image sequences was performed in MATLAB (version R2014a, The MathWorks Inc., Natick, Massachusetts).

2.5.2 Computational analysis

Computational analyses were developed with and carried out by Dr Leila Muresan (Cambridge Advanced Imaging Centre). Analyses of 3D image stacks, selected on the basis of their resolution and compatibility with 3D image analysis, aimed to identify ducts within the intact mammary stroma and to subsequently recognize all ductal YFP^+ cells. Ductal YFP^+ cells were classified as luminal or basal based on the proportion of K8 versus SMA fluorescence signal.

For a volumetric analysis, the volume ratio of YFP^+ cells within each duct was computed with respect to the entire ductal (cellular) volume, and the intensity of K8 in YFP^+ cells was also compared with the overall K8 intensity level within the duct. For computational efficiency, a multi-resolution transform was used for the K8 channel; a coarse scale was used to detect the duct and the full detail scale was used to identify voxels significantly different from the background. The coarse scale was segmented with a robust threshold, obtained as the median of the intensity values in the transformed stack plus three times the median absolute deviation of these values. The up-sampled structure represents an approximation of the duct and the sum of all its voxels was a measure of the volume of the duct.

Within the duct, significant voxels (excluding intercellular spaces and nuclei) were detected plane-wise from the fine detail levels of the wavelet coefficients of the two-dimensional wavelet transformed image by applying a false discovery rate-based thresholding. Independently, YFP^+

cells were identified after a difference of Gaussian filtering suited to the noise level of the image; as the filter was applied on the full-resolution 3D stack, a recursive filter implementation in CImg (Deriche, CImg) was used for time efficiency. The threshold was computed as above, as the robustly estimated 99% quantile of the Gaussian distribution of filtered intensity values: the median plus three times the median absolute deviation of these intensities. Subsequently, only YFP^+ cells inside the detected duct were taken into account. These cells were classified as luminal or basal based on the comparison of intensity values in the K8 and SMA channels of the voxels belonging to each segmented cell; for a chosen threshold, the voxels exceeding this threshold in the K8 and SMA channels, respectively, were counted. If the number of K8 voxels exceeded the SMA, the cell was classified as luminal, otherwise it was classified as basal. Note that a perfect exclusion of one colour cannot be expected due to the resolution limits of the optical system.

To provide robustness with respect to the choice of the intensity threshold, the classification was performed using a multi-threshold approach, with levels 100, 300, 500, up to 1,500 and the majority vote for all thresholds gave the final classification of the cell. Finally, a Kolmogorov–Smirnov test was performed to determine if the significant voxels of the duct in K8 channel were differentially distributed compared with the significant voxels inside the segmented YFP^+ cell.

Some particularities of the basal YFP^+ co-localization images (such as the poorer signal-to-noise ratio of these images due to the depth of this clone within the mammary fat pad and the elongated shape of the YFP^+ cells) meant modifications of the described analysis was necessary. To improve the quality of the images, a denoising step was applied followed by a fast deblurring step (Dr Jerome Boulanger, Medical Research Council Laboratory of Molecular Biology, private communication). The segmentations are performed in 3D for all channels, however, to separate elongated and overlapping YFP^+ cells, a seeded watershed was used (the seeds are thresholded distance images of the inverted segmented YFP^+ image, where the threshold is manually selected).

The classification of the YFP^+ cells was performed as before: if the number of K8 voxels exceeded the SMA, the cell was classified as luminal, otherwise it was classified as basal. Using this analysis, two cells were excluded from the classification due to their localization in regions

where the SMA signal was undetectable and thus the double/nested tubular structure could not be observed.

2.5.3 PR / K8 association analysis

Quantification of PR^+ and K8hi cells was performed on maximum intensity projections of 3D image stacks using the Cell Counter plugin in Image J (v1.50a, National Institutes of Health). Maximum intensity projections of PR and K8 channels were scored independently. At least 200 cells were counted per image, with six images analysed from three independent mice (total cells counted: 1,831). Manual counting of YFP^+ cells was performed on 10 histological sections cut 425 mm apart, spanning 300 mm. K8 was used to mark the luminal lineage and SMA was used to mark basal cells. The number of alveoli that were fully or partially populated by YFP cells of a single lineage were manually counted in Image J.

2.5.4 Nearest neighbour clonal analysis

A cell neighbour analysis was used to analyse labelling outcomes in the R26R-Confetti model in both embryonic and pubertal labelling. Quantification method was based off previous pair/patch-based analysis (Wuidart et al., 2018), with manual scoring (Van Keymeulen et al., 2011; van Amerongen et al., 2012; Rios et al., 2014). 3D image stacks of label positive regions were randomly selected based on image quality (and therefore ability to accurately record cell lineage via cell morphology, topology and SMA positivity). For all cells within a region, the lineage of its closest same-colour neighbour was recorded in one of two bins: "same" or "different". For example, if a luminal YFP+ cell was observed and its closest YFP+ neighbour (in x-y-z dimensions) was also luminal, this cell would be counted as "same", and thus potentially supporting unipotency of the clone. On the other hand, if its closest YFP+ neighbour was basal, this would be counted as "different" and thus potentially supporting bipotency of the clone.

Hundreds of cells were counted over different clones in different mice to obtain a cell neighbour lineage analysis. Based on the relatively low level of labelling and the multi-colour reporter system, this method of analysis assumes that two proximal cells of the same colour within a specific region were derived from a common labelled ancestor. Thus, two luminal YFP+ cells counted as "same" may be distinct labelling events derived from different precursor cells – as such, this model may over-represent the percentage of possible unipotent clones. Similarly, a luminal and basal YFP+ cell counted as "different" may also be from distinct labelling

events and thus this model may also overrepresent the proportion of possible bipotent clones, which was extremely small in this study.

2.6 Statistics.

To determine if a single labelled mammary stem/progenitor cell contributed equally to both the $K8^{hi}/PR^{+}$ and $K8^{lo}/PR^{-}$ luminal populations, the Kolmogorov–Smirnov test was performed on custom 3D imaging algorithms in MATLAB (R2014a, The MathWorks Inc., Natick, Massachusetts). Briefly, this test was performed in order to determine if the significant voxels of the segmented total ductal luminal cells were differently distributed to the significant voxels inside segmented YFP^{+} cells. All values are shown as mean±s.d.

Chapter 3

Development of optical clearing and 3D imaging techniques for use in the mammary gland

3.1 Introduction

3.1.1 Imaging requirements in the mammary gland

As described previously, the mammary gland is a complex, branched organ embedded in a fatty stroma. Numerous primary and secondary ducts occupy multiple focal planes which are remodelled during oestrous cycling and with gestation/lactation/involution cycles (Watson and Khaled, 2008). As such, this can present issues when visualising the mammary gland. Like other tissues, traditionally the gland has been imaged using 2D sections of either paraffin-embedded or frozen tissue. Taking multiple slices, usually only approximately 10μM thick, allows antibody-based immunohistochemical staining which can result in excellent cellular resolution. However, cells and tissues are inherently 3D, and thus analysing tissues in 2D can present issues when placing the section in the context of both other branches, as well as the entire gland.

This becomes particularly relevant when interpreting lineage tracing data, whereby the progeny of labelled stem or progenitor cells are analysed in order to deduce characteristics of the parent cells – such as lineage, location and proliferation capacity. Traditional tissue sections can be insufficient for accurate quantitative analysis and do not reveal complex spatial patterns that may be present. As such, visualising the mammary gland in 3D, with the entire ductal tree intact, is an important step in correctly describing mammary gland development, and its stem and progenitor cells and their progeny.

3.1.2 Optical clearing: theory and development

Other tissues with branching characteristics have presented similar difficulties. This, combined with advancing imaging techniques which have allowed increased imaging depth, has necessitated the development of methods that allow for imaging whole tissues in three dimensions. As such, tissue optical clearing techniques have been developed, which render tissue transparent to allow for imaging entire tissues in one piece.

The general theory of optical clearing involves using different techniques to homogenise the refractive indices of tissues. Different cellular and intracellular components have differing refractive indices depending on the contents and density of its constituents, such as proteins and lipids (Johnsen and Widder, 1999; Tuchin, 2015). When light travels through structures with differing refractive indices it is bent, or refracted, causing spherical aberrations, resulting

in the inability to image any deeper. Moreover, this scattering of light results in the loss of excitation and emission efficiency, resulting in a lower resolution and imaging depth when using fluorescence microscopy.

Optical clearing techniques have been largely spearheaded by the neuroscience field, with the development of CLARITY for whole brain imaging in 2013(Chung et al., 2013). Protocols can be generally divided into active or passive: active optical clearing involves the use of electrophoresis chambers that, over time, perfuse the tissue with the clearing solution. Passive clearing, on the other hand, simply involves immersion of tissues into clearing solutions, and as such less equipment and set up is required, meaning they are generally more economical.

Despite the peak in the development of optical clearing techniques within the last decade, tissue clearing is not a new concept. German anatomist Werner Spalteholz first described the use of "aufhellungspräparate", or brightening preparations, in the 19th century, whereby the coronary arteries of the heart were imaged following impregnation with methyl salicylate, benzyl benzoate and wintergreen oil by immersion in a solution of these reagents in order to obtain matched refractive indices (Spalteholz, 1911), in a similar manner to modern-day protocols. Recent optical clearing methods are based on the same concept of refractive index homogenisation, but with a shift towards more sophisticated approaches that utilise more current techniques, which allows tissue integrity to be maintained whilst also requiring much shorter protocol times.

The recent renaissance of optical clearing has also been brought about by advancement in imaging methods that are able to image whole tissues. There have been significant advances in the microscope since the development of the modern day model by Antonie van Leeuwenhoek in the 17th century (Azaripour et al., 2016), with the introduction of the ability for optical sectioning. Confocal microscopy is now a standard tool for scientists, allowing imaging of multiple colours sequentially. Moreover, recent developments have led to light sheet microscopy also gaining popularity.

Although 3D imaging of the mammary gland has been done previously, the protocols used result in a number of limitations. This has been achieved largely using either microdissection or proteolytic digestion, although both have a number of caveats. Microdissection involves removing parts of the epithelial ductal tree from its native stroma (Rios et al., 2014), therefore

negating the need to image through the large volumes of adipocytes that are present in the gland, and are especially light refractive (Rios et al., 2014). However, this presents a number of issues. Firstly, this often means that only small sections of the ductal network can be imaged, in addition for the need of a high degree of technical skill. Secondly, it rules out using 3D imaging in the study of the native stromal compartment of the mammary gland.

The other method previously used for 3D imaging of the mammary gland involves a combination of enzymes, such as collagenase and hyaluronidase, to dissolve the extracellular matrix – and in the case of the mammary gland – the surrounding adipocytes (Wuidart et al., 2016). However, achieving specific digestion of the stroma whilst also leaving the epithelium intact is a delicate balance. Consequently, recent work has shown that the use of proteolytic digestion for 3D imaging in the mammary gland has a detrimental effect on the epithelium (Rios et al., 2016). The basal cell layer was shown to be particularly affected, presumably due to the destruction of the basal lamina, which basal cells rely on for cell contact. Using even low concentrations of proteolytic enzymes resulted in depletion of cells in the outermost layer of the ductal tree, leading to gaps. The remaining cells displayed altered morphology, losing their traditional elongated striated shape and becoming more rounded (Rios et al., 2016).

As such, optical tissue clearing provides a method to yield the benefits of imaging the entire mammary gland in 3D, without detrimental effects on either the stroma or the epithelium. Imaging the ductal network, *in situ* and without alteration, removes potential complications or anomalies that arise and that may contribute to disparities in the interpretation of data in the field of mammary gland research.

3.1.3 Aims

The current imaging methods in the mammary gland may lead to the inability to properly interpret the data, and as such may be contributing to many of the disparities that are currently affecting mammary gland research. As such, we aimed to develop and optimise optical tissue clearing methods for use in the mammary gland, based on the adaptation of published protocols, to allow wholemount imaging. Moreover, we aimed to assess multiple imaging methods in order to determine microscopy techniques that can be combined with optical tissue clearing for different uses.

To assess the clearing methods in the mammary gland, in the first instance, methods were evaluated based on the ability to render the gland macroscopically transparent. However, the compatibility with immunohistochemical staining, including wholemount colourimetric and antibody-based immunofluorescence, were also considered as of primary importance, as well as maintenance of tissue integrity.

3.2 Results

3.2.1 Assessment and selection of optical clearing protocols

There are multiple optical clearing protocols that have been developed, and as such this required a strategy to narrow down options. To begin, we decided to focus on passive optical clearing due to both the relatively gentler treatment of the tissue (in comparison to active optical clearing, involving often harsh electrophoresis). Additionally, generally passive optical clearing protocols provide a more economical option – requiring both less cost and time – as well as being more user-friendly due to not requiring specialist equipment. Due to the fact that a number of tissues would be analysed in future projects, these criteria were considered to be important.

The mammary gland is a relatively delicate epithelial organ, but with a high level of fat content. Consequently, a fine balance would need to be considered in order to achieve an adequate level of clearing, whilst still maintaining tissue integrity, such as morphology - both on the larger scale in terms of the ductal network, and on the cellular scale in terms of cell-cell contact. For imaging of optically cleared tissue, we assessed the protocols in the first instance for compatibility with confocal microscopy. The majority of laboratories have access to a confocal microscope, and is the most available method in our laboratory. Moreover, with its ability to optically section, it is reasonably easy to adapt to imaging of wholemount tissues.

3.2.1.1 SeeDB optical clearing

The first clearing protocol we elected to examine was the water-based "See Deep Brain" (SeeDB). Developed to allow reconstruction of neuronal circuits, SeeDB clearing utilises increasing concentrations of fructose to match the refractive indices of tissues, combined with α-thioglycerol to prevent Maillard reaction-based tissue browning (Dills, 1993; Ke et al., 2013), resulting in a total protocol length of 15 days, including antibody staining (Figure 3.2.1.1a). This protocol aims to overcome the limitations of previously published protocols - namely long incubation times and tissue deformation. The gentle nature, with the absence of harsh detergents and other chemicals, also means that theoretically tissue integrity is highly preserved – for example in terms of epitopes for antibody binding or fine tissue structures. This makes SeeDB particularly amenable for multicolour fluorescent-based immunohistochemistry, particularly when combined with endogenous fluorescent protein expression.

When tested on the mammary gland, SeeDB optical clearing lead to adequate optical transparency in the virgin mammary gland (Figure 3.2.1.1b, left). An increase in optical transparency was also achieved in the lactating mammary (Figure 3.2.1.1b, right), but to a lesser degree than in the virgin – presumably due to the increase in epithelial density associated with lactation.

Although SeeDB did not provide the highest degree of optical transparency compared to other protocols (judged by the degree of visibility of the reference grid), it was however adequate enough for use with confocal imaging (Figure 3.2.1.1c), with both luminal and basal cells visualised using K8, E-cad, K5 and SMA antibodies. Importantly, due to the inert nature of fructose, most antibodies tested were highly compatible with this method.

The longer incubation times required in the SeeDB protocol have the potential to cause browning of the tissue via the Maillard reaction, due to the interaction between the amino acids and proteins in the tissue and the fructose in SeeDB (Dills, 1993). The result of this is the formation of both brown and fluorescent by-products, which can lead to high levels of background fluorescence accumulation. Moreover, the loss of amino acid residues as a result of the reaction results in a reduction in protein quality, which could lead to a decrease in antibody binding due to changes in epitope structure (Dills, 1993). However, in the mammary gland, there was a very low level of autofluorescence background (Figure 3.2.1.1d), and the protocol was found to be compatible with a number of antibodies, presumably due to the inclusion of α-thioglycerol in the fructose solutions which inhibits the Maillard reaction. Consequently, SeeDB was assessed as a compatible method for use with immunofluorescence-based confocal imaging, especially in situations whereby new or sensitive antibodies are to be used.

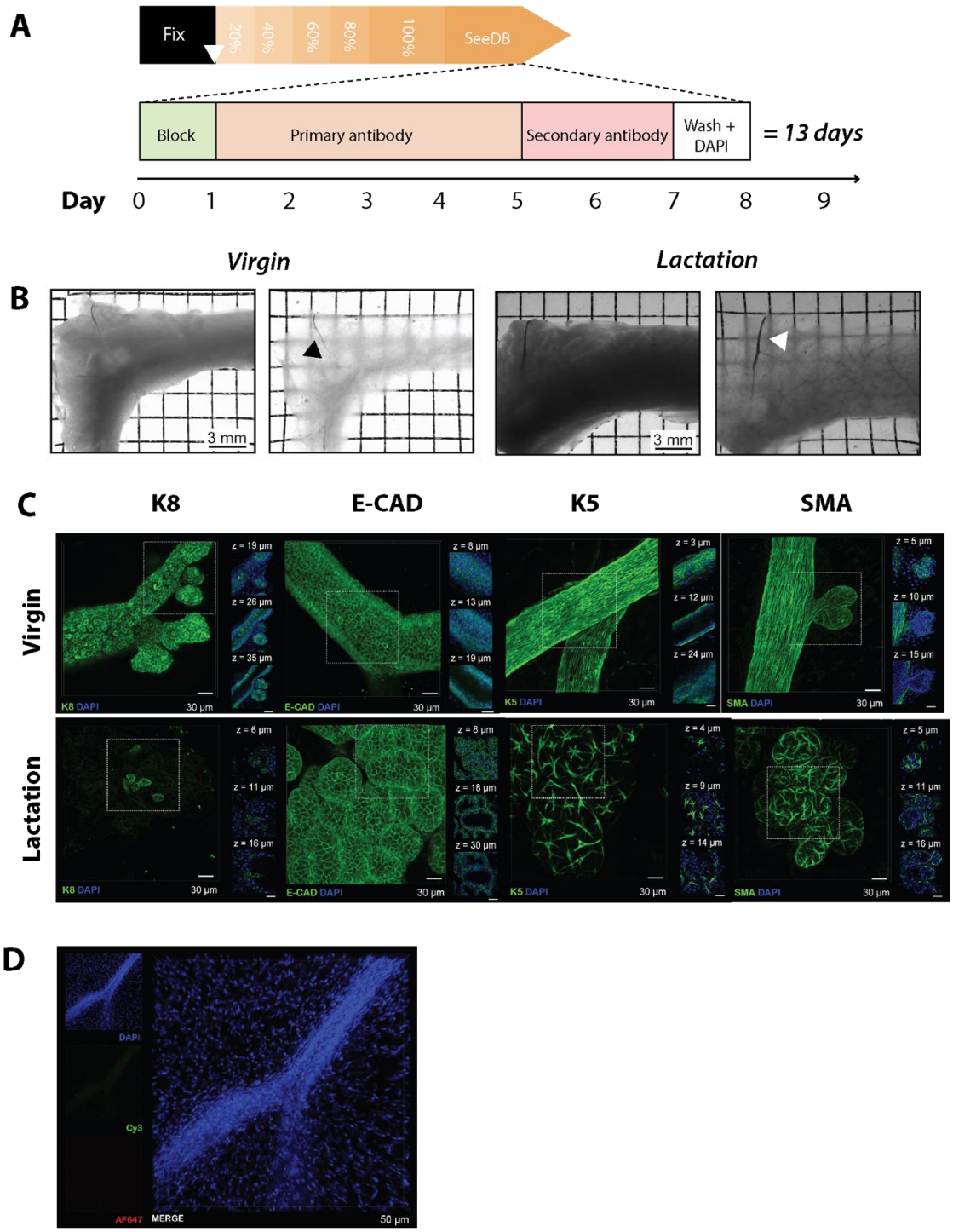


Figure 3.2.1.1 SeeDB-based optical tissue clearing in the mammary gland

(A) Schematic outline of the SeeDB optical clearing protocol, including primary and secondary antibody staining, resulting in a total protocol time of 13 days. **(B)** Use of SeeDB in both virgin and lactating tissue resulted in adequate transparency, without any decolourisation (shown by the presence of blood vessels, arrow heads, after optical clearing. **(C)** Fluorescent antibody-based staining using K8 and E-CAD for luminal cells, and K5 and SMA for basal cells was used on virgin and lactating mammary gland tissues (top and bottom panels, respectively). **(D)** No-primary control tissue, with only secondary antibodies used, showed a very low level of background and autofluorescence on mammary tissue with this SeeDB protocol.

3.2.1.2 CUBIC optical clearing

The second immersion-based method that was tested is known as CUBIC (Clear, Unobstructed Brain Imaging Cocktails) (Susaki et al., 2014). CUBIC clearing is composed of sequential incubations in the two reagents, reagent 1 (R1) and regent 2 (R2). R1 is designed to dissolve fat and remove pigment, being composed of high concentrations of detergents, such as urea and Triton-X100. R2 also contains triton X-100 and urea but at lower concentrations that R1. The inclusion of a high level of sucrose in R2 lends to its primary utility in matching the refractive index of tissues.

In this clearing protocol, we found incubation in R1 and R2 for 72 and 48 hours respectively, was optimal for the mammary gland, resulting in a total protocol time of 15 days when including antibody staining (Figure 3.2.1.2a). The higher concentration of detergents in this protocol lead to a higher degree of transparency in comparison to SeeDB, with the additional benefit of removal of pigment - which is seen by the disappearance of blood vessels, due to the discolouration of haem (Figure 3.2.1.2b, arrowheads). Strong immunostaining of basal cells was observed using antibodies such as K5 and SMA (Figure **3.2.1.2**c, right). However, in some cases, the use of stronger detergents also came with an associated decrease in antibody viability: immunostaining with antibodies such as K8 and E-cad resulted in non-uniform expression, presumably due to changes in epitope structure induced by the detergents themselves (Figure 3.2.1.2c, left).

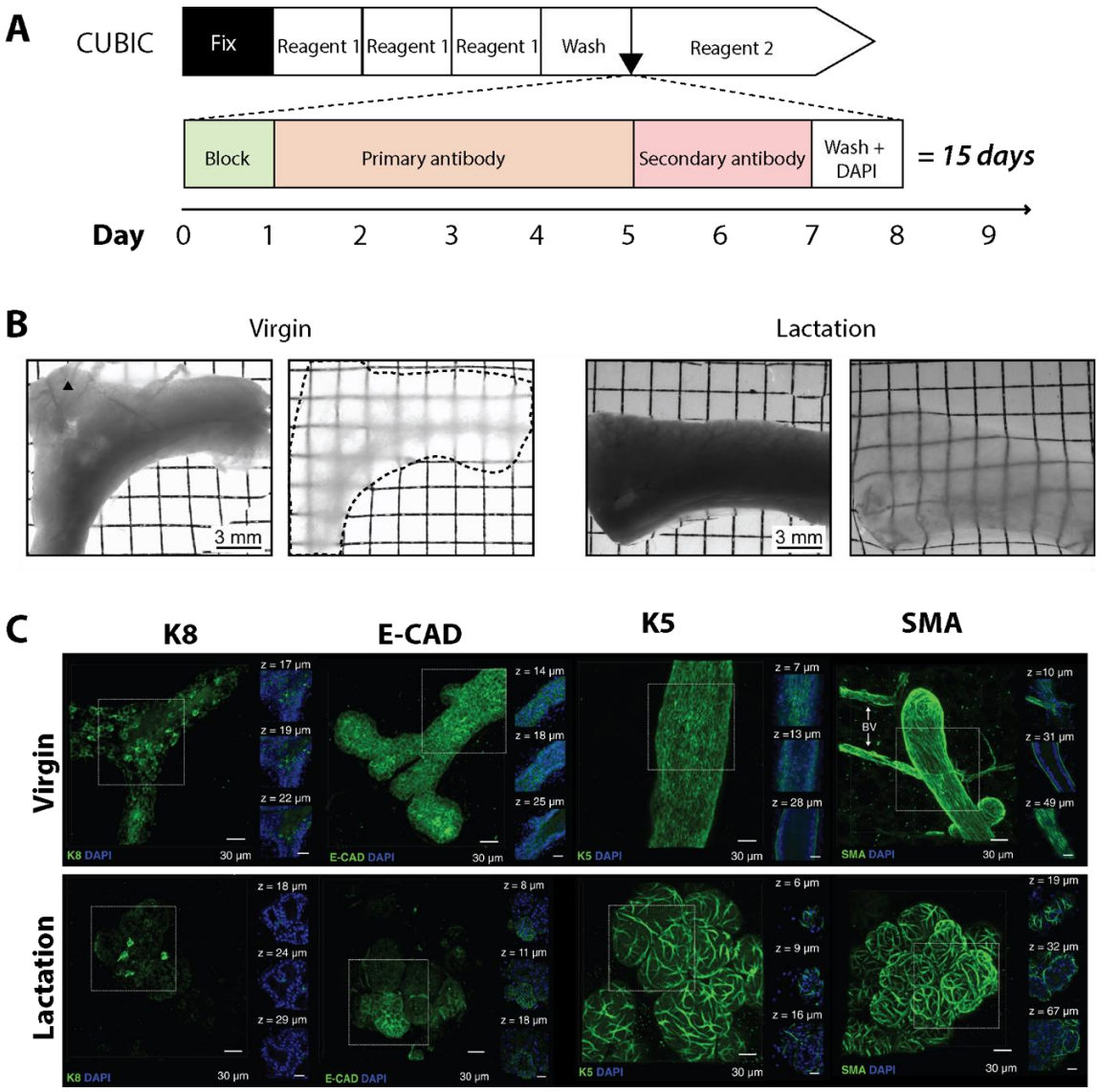


Figure 3.2.1.2 CUBIC-based optical tissue clearing in the mammary gland

(A) Schematic outline of the CUBIC optical clearing protocol, including primary and secondary antibody staining, resulting in a total protocol time of 15 days. **(B)** Use of CUBIC in both virgin and lactating tissue resulted in excellent transparency, with the additional benefit of decolourisation (shown by the absence of blood vessels after optical clearing (arrow head). **(C)** Fluorescent antibody-based staining was used in both virgin and lactating mammary gland tissues (top and bottom panels, respectively). Using K8 and E-Cad, for luminal cells, showed immunostaining was not as compatible with CUBIC clearing as other protocols. However, K5 and SMA staining for basal cells worked well. *BV*, blood vessels.

However, a recent modification of the CUBIC protocol, termed CUBIC R1a (Susaki and Ueda, 2016), has been developed as a more gentle alternative. CUBIC R1a uses lower concentrations of detergents, with a third less triton-X and a two-thirds reduction in urea. Using this modified formulation, we were able to observe improved immunostaining of both K8 and E-cad (Figure 3.2.1.3a) - suggesting that the higher concentration of detergents in the original CUBIC R1 formulation does indeed affect the preservation of epitope availability. However, in both the new and original protocols, levels of autofluorescence were observed within some ducts and blood vessels, which could be seen on both confocal (Figure 3.2.1.3b) and stereo microscopy (Figure 3.2.1.3c). Significant background was also seen on the no primary control (Figure 3.2.1.3d). This could be due to the accumulation of the clearing solution within closed off structures, such as blunt-ended ducts, and as such, we decided to increase washing times to try and prevent this from occurring.

Due to the excellent level of transparency achieved and relative ease of the protocol, we also opted to assess the use with wholemount colourimetric stains, which require high transparency for the ability the see through the entirety of the tissue. CUBIC-based clearing was compatible with the enzymatic reaction used to detect β-glucosidase (SYNbglA), a magenta coloured reporter similar to LacZ (McCutcheon et al., 2010; Kozar et al., 2013) (Figure 3.2.1.4a). Due to the magenta colour of the SYNbglA reporter being too similar to the traditional carmine stain used in the mammary gland (Figure 3.2.1.4b), we also needed to use a different coloured epithelial co-stain. For this, we opted to test methyl green, a traditional DNA counterstain used in IHC (Høyer et al.). This worked well to show the epithelial compartment of the mammary gland, with the entire ductal network easily visualised along with a low background (Figure 3.2.1.4c).

We also utilised diaminobenzidine (DAB) staining with CUBIC clearing, an often-used enzyme-based chromogenic stain in 2D immunohistochemical analysis (Singer, 1959). We found using the original CUBIC R1 protocol and DAB staining combined with SMA, either after R1a or R2, lead to good signal, especially at high resolution (Figure 3.2.1.4d). This suggests that CUBIC clearing does not disrupt the ability to use a number of enzymatic-based counterstains. We therefore adopted use of CUBIC for colourimetric staining, where the importance of visualising the whole tissue in transmission imaging outweighs the need for specific antibody binding.

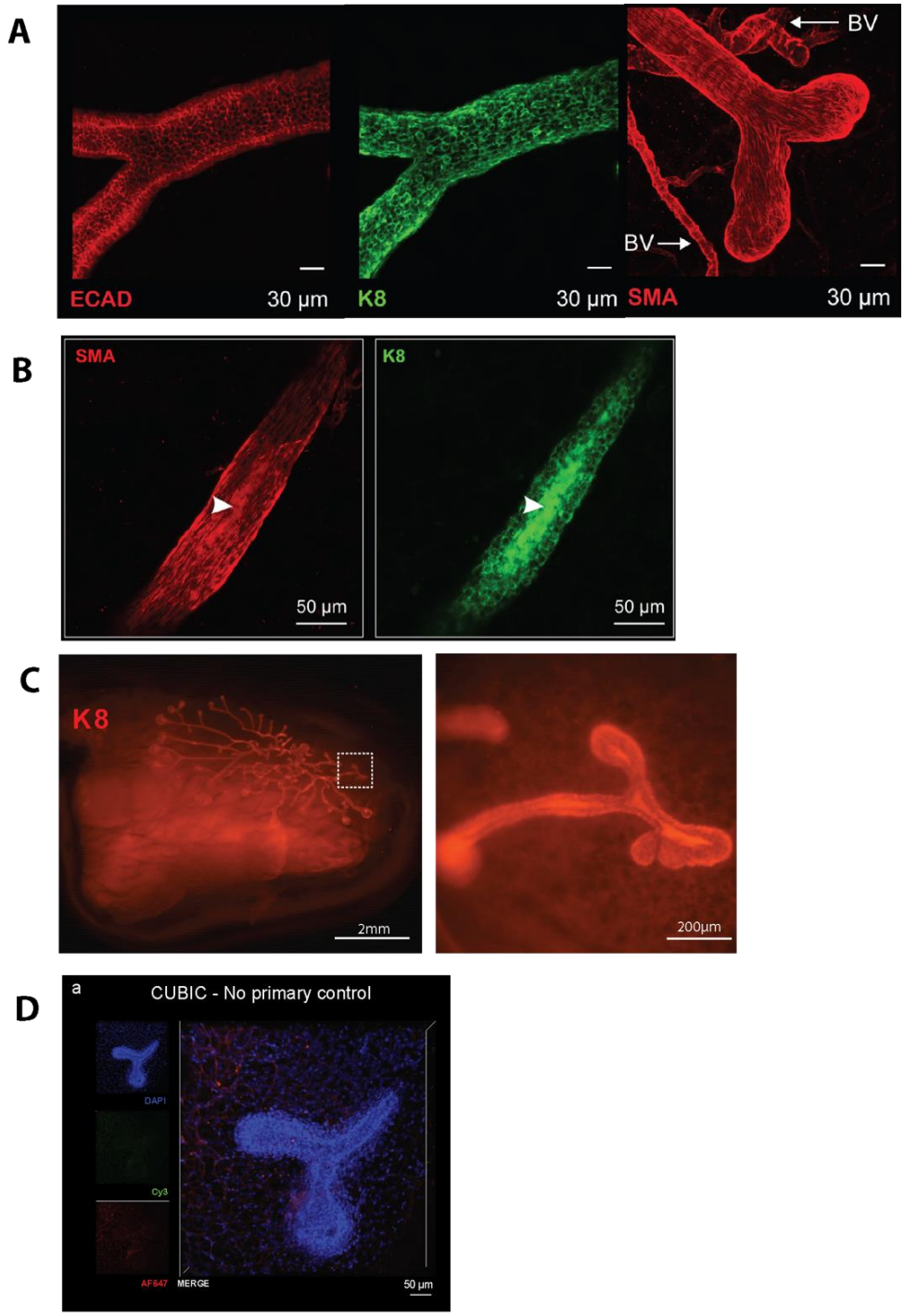


Figure 3.2.1.1 CUBIC R1a-based optical tissue clearing in the mammary gland

(A) CUBIC R1a clearing was more compatible than the original CUBIC clearing cocktail for subsequent fluorescent antibody-based staining. Images show examples of staining for luminal cells using K8 and E-Cad and for basal cells using SMA in virgin mammary gland. **(B)** However, both CUBIC protocols resulted in high levels of non-specific background fluorescence within ductal structures when visualised using confocal microscopy (arrow heads). **(C)** This could also be detected when visualised using stereo microscopy (example image shows 4 week mammary gland) **(D)** A background autofluorescent signal could also be seen in the no primary control, with secondary antibody staining using Cy3 and AlexaFluor647. *BV,* blood vessels

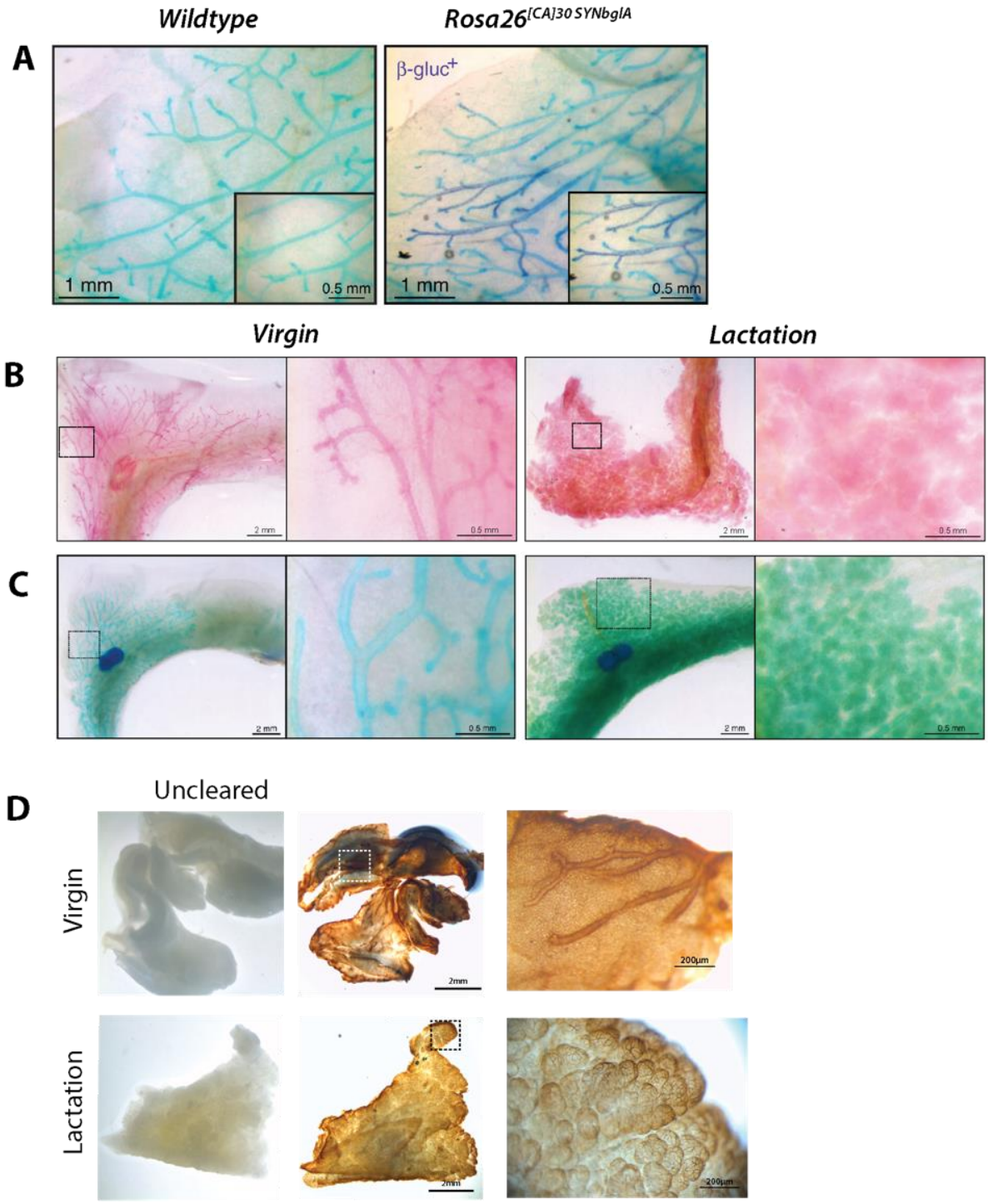


Figure 3.2.1.2 CUBIC-based optical tissue clearing combined with colourimetric staining in the mammary gland

(A) CUBIC clearing was compatible with the enzymatic reaction required to stain for the SYNbglA reporter **(B)** CUBIC clearing was also compatible with carmine staining, in addition to methyl green counterstain, as shown in **(C)**. **(D)** CUBIC clearing could also be combined with DAB immunohistochemistry, with example images showing SMA staining in both virgin (top) and lactating tissue (bottom).

3.2.1.3 PACT optical clearing

We next assessed a hydrogel-embedding based clearing protocol, which is designed to provide extra sub-cellular stabilisation of tissue integrity due to the protein cross-linking (Chung et al., 2013). PAssive CLARITY Technique (PACT)-based tissue clearing involves three main steps: first, stabilisation of biomacromolecules is achieved using hydrogel embedding to facilitate crosslinking and hybridization to hydrogel monomers (Yang et al., 2014). Next, ionic detergents are used to extract lipids from the issue, before a final third step of embedding in a refractive index matching solution (RIMS) for refractive index homogenisation for either imaging or long-term storage (summarised in Figure 3.2.1.5a).

PACT-RapiClear uses a commercially available RIMS as the final step. When used in both virgin and lactating tissue, excellent transparency was achieved (Figure 3.2.1.5b). However, when we imaged using confocal microscopy after immunostaining with SMA, the signal was patchy and had a high background level, especially in virgin tissue (Figure 3.2.1.5c, top panels). Moreover, RapiClear is used with specialised imaging chambers for final embedding to allow for longer term storage.

However, when used with both virgin and lactating mammary gland samples, imaging was restricted due to the limited working distance of most commercial confocal microscope objectives. We found that using the imaging chambers meant the tissues could not be properly orientated against the cover glass, meaning optimal sample illumination and imaging was difficult to achieve, especially with thicker samples, such as those in the lactating mammary gland (Figure 3.2.1.5c, bottom panel). This, combined with the commercial aspect of RapiClear, which makes it much less economical than other clearing agents, meant we sought to find an alternative RIMS to use with PACT clearing.

As such, we opted to test PACT with the non-commercial refractive index-matching solutions. The first, RIMS, uses a high concentration of a solution of Histodenz, a chemical often used as a radiocontrast agent (Ye et al., 2014). The concentration can be adjusted depending on the RI of the tissue in question, up to an 88% w/v solution which corresponds to a refractive index of 1.48 (Yang et al., 2014). The refractive index of the murine mammary gland has not been tested, and as such estimations based on similar tissues were
made. The refractive index of adipose is generally one of the highest of all tissue types (Susaki et al., 2014), with bovine adipose tissue samples being previously measured at 1.455 (Bolin et

al., 1989). One study in rat mammary tissue has corroborated this – with rat mammary adipose and stroma being measured at 1.467 and 1.388 respectively (Zysk et al., 2006). As such, we decided to test 20 – 88% w/v histodenz, which encompasses refractive indices from 1.38 – 1.48 and thus encompasses the ranges likely to be found in murine mammary tissue.

After PACT hydrogel embedding in A4P0 and clearing in sodium dodecyl sulfate (SDS), virgin mammary gland tissue fragments were incubated in increasing concentrations of histodenz (20-88% w/v), and the transparency levels over the next few hours observed. PACT clearing resulted in pigment removal (as shown by the absence of blood vessels) but no significant changes in transparency or change in tissue volume (Figure 3.2.1.6a). Significant changes in the tissue transparency were not apparent until concentrations above 60%, presumably due to the high RI of the fat in the mammary gland - suggesting that the RI of the murine mammary gland is above 1.44. Supporting this, incubation in 88% histodenz resulted in noticeable transparency changes after only 2 hours, with a maximum level seen after 48 hours. However, we found RIMS to be both difficult to dissolve, and associated with a high cost. We therefore decided to asses next the alternative, sorbitol refractive index matching (sRIMS), which utilises a sorbitol-based solution for RI matching.

Following PACT clearing, transparency was seen after two hours, similar to RIMS. After 48 hours, maximum transparency was achieved, but to a slightly lesser extent than with RIMS. Due to sorbitol being a common laboratory reagent, sRIMS has the benefit over both histodenz-based RIMS and RapiClear in that it is both inexpensive and easier to source (at approximately £0.11 per millilitre it is over 40 times more economical than RIMS).

We therefore decided to combine PACT-sRIMS and immunofluorescent staining with confocal imaging. In both the virgin and lactating tissue, we were able to image the mammary gland using the standard antibody panel (Figure 3.2.1.6b). However, in some cases, there appeared a high degree of background, which can complicate downstream computational image analysis. However, this was improved with the application of 3D de-noising algorithms (Boulanger et al., 2010), which although not overtly affecting the outward appearance of the image stacks, would be beneficial for any subsequent computer-assisted analyses. This, combined with the longest total protocol time of 20 days, meant PACT-sRIMS did not provide any benefit over other protocols for day to day optical tissue clearing in the mammary gland when combined with confocal microscopy.

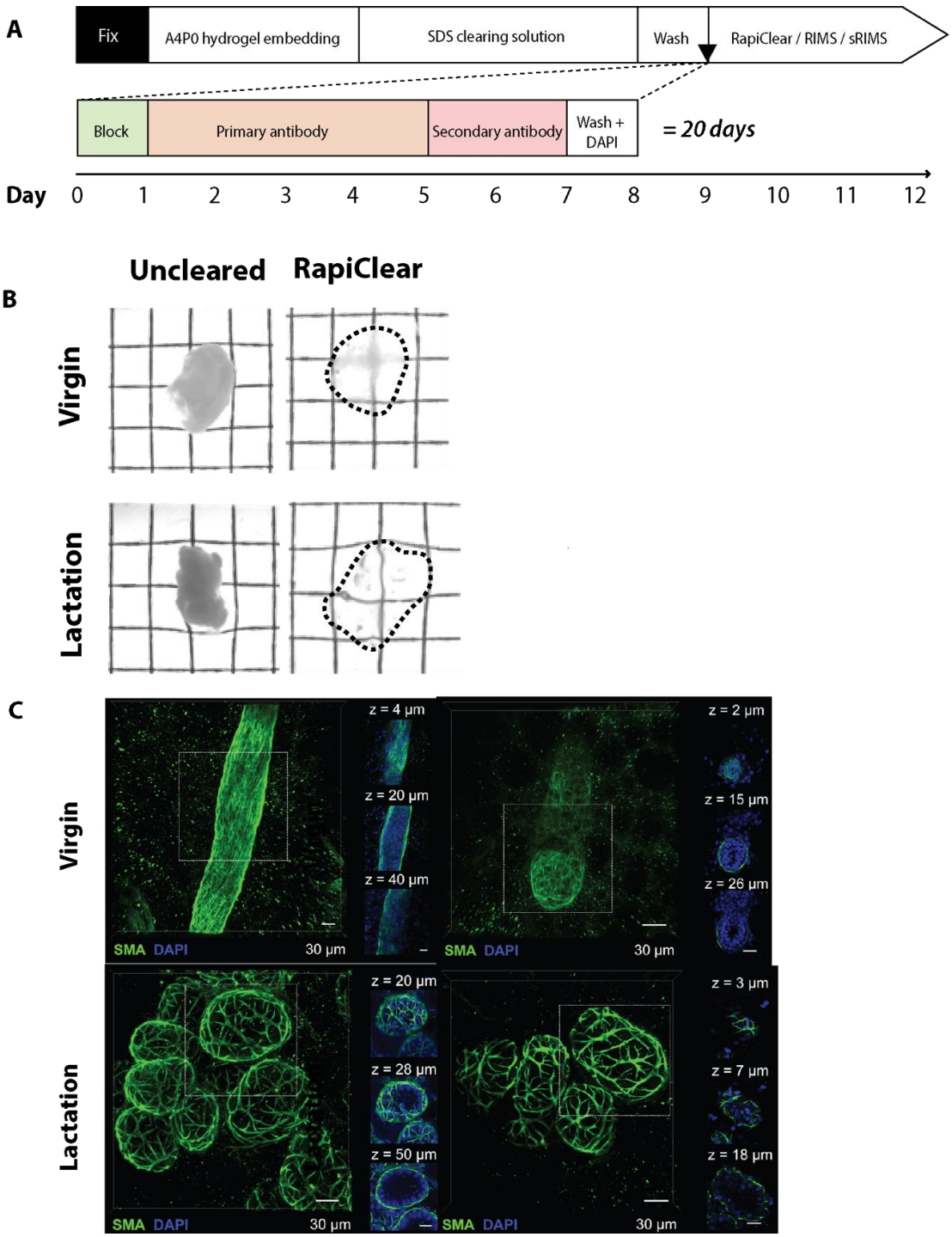


Figure 3.2.1.3 PACT-RapiClear optical tissue clearing in the mammary gland

(A) Schematic outline of PACT-RapiClear optical clearing protocol, including primary and secondary antibody staining, resulting in a total protocol time of 20 days. **(B)** Fragments of virgin mammary tissue were cleared with the PACT protocol. Tissue was then incubated in RapiClear for refractive index matching. **(C)** PACT clearing and RapiClear refractive index matching were used in both virgin and lactating mammary gland (top and bottom panels, respectively). This was combined with fluorescent antibody-based staining for SMA and imaged using confocal microscopy.

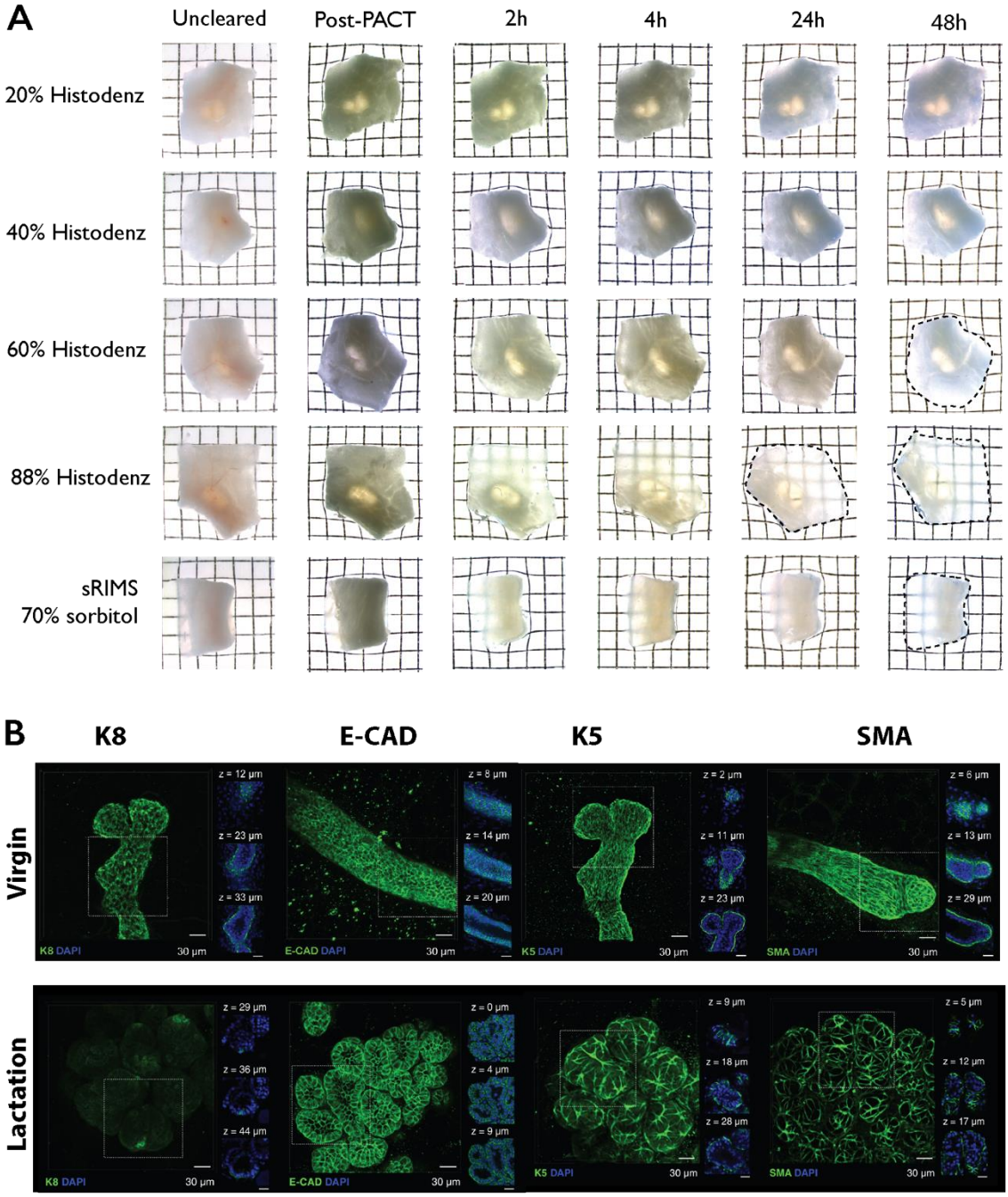


Figure 3.2.1.4 PACT-RIMS and sRIMS optical tissue clearing in the mammary gland

(A) Fragments of virgin mammary tissue were cleared with the PACT protocol (second panel). Tissue was then incubated in increasing concentrations of histodenz (20-80%), or 70% sorbitol (bottom row) for refractive index matching. **(B)** PACT clearing and sRIMS refractive index matching were used in both virgin and lactating mammary gland (top and bottom panels, respectively). This was combined with fluorescent antibody-based staining. Using K8 and E-Cad, for luminal cells, and K5 and SMA for basal, cells allowed for good visualisation at single cell resolution using confocal microscopy.

3.2.1.4 3DISCO optical clearing

Solvent-based tissue clearing of the mammary gland using "3D imaging of solvent-cleared organs" (3DISCO) was previously described in the original protocol (Ertürk et al., 2012). This protocol consists of three steps: initial dehydration with increasing concentrations of the solvent tetrahydrofuran (THF), lipid extraction with incubation in dichloromethane (DCM) and RI matching with immersion in dibenzyl ether (DBE) (summarised in Figure **3.2.1.**7a). 3DISCO is well suited to lipid-rich tissues due to the de-lipidation steps, and so in theory should be well-matched for use in the mammary gland.

We used 3DISCO clearing in both virgin and lactating mammary tissue, which resulted in very good transparency in comparison to other immersion-based clearing methods (Figure 3.2.1.7b). Moreover, it involved fewer steps, and as such less time, with a total protocol length of only 12 days (Figure 7a). Despite this, we found its utility in the mammary gland to be limited for a number of reasons. Despite the shorter incubation times, the mammary gland was left brittle and showed a significant level of damage. On the macro level, this resulted in significant shrinkage in both virgin and lactating tissues (Figure 3.2.1.7c). This was also reflected on the microscopic level, with reduced staining of SMA, in addition to the virgin ducts and lactational alveoli displaying structural deformations when imaged using confocal microscopy (Figure 3.2.1.7d)

Additionally, on a practical level, the solvents utilised in 3DISCO pose a number of safety risks and as such require specialist imaging chambers, equipment and disposal of imaging solutions (for example, THF negates the use of any plastic). This also means that 3DISCO is incompatible with most immersion objectives. Moreover, the use of these strong solvents negates the ability to store tissue for prolonged periods of time, particularly in samples with endogenous fluorescence. The authors of the original manuscript also stated that less stable fluorescent proteins, such as YFP, are particularly affected. For these combined reasons, although 3DISCO may have utility in other tissues and organs, it was not pursued further for use in the mammary gland.

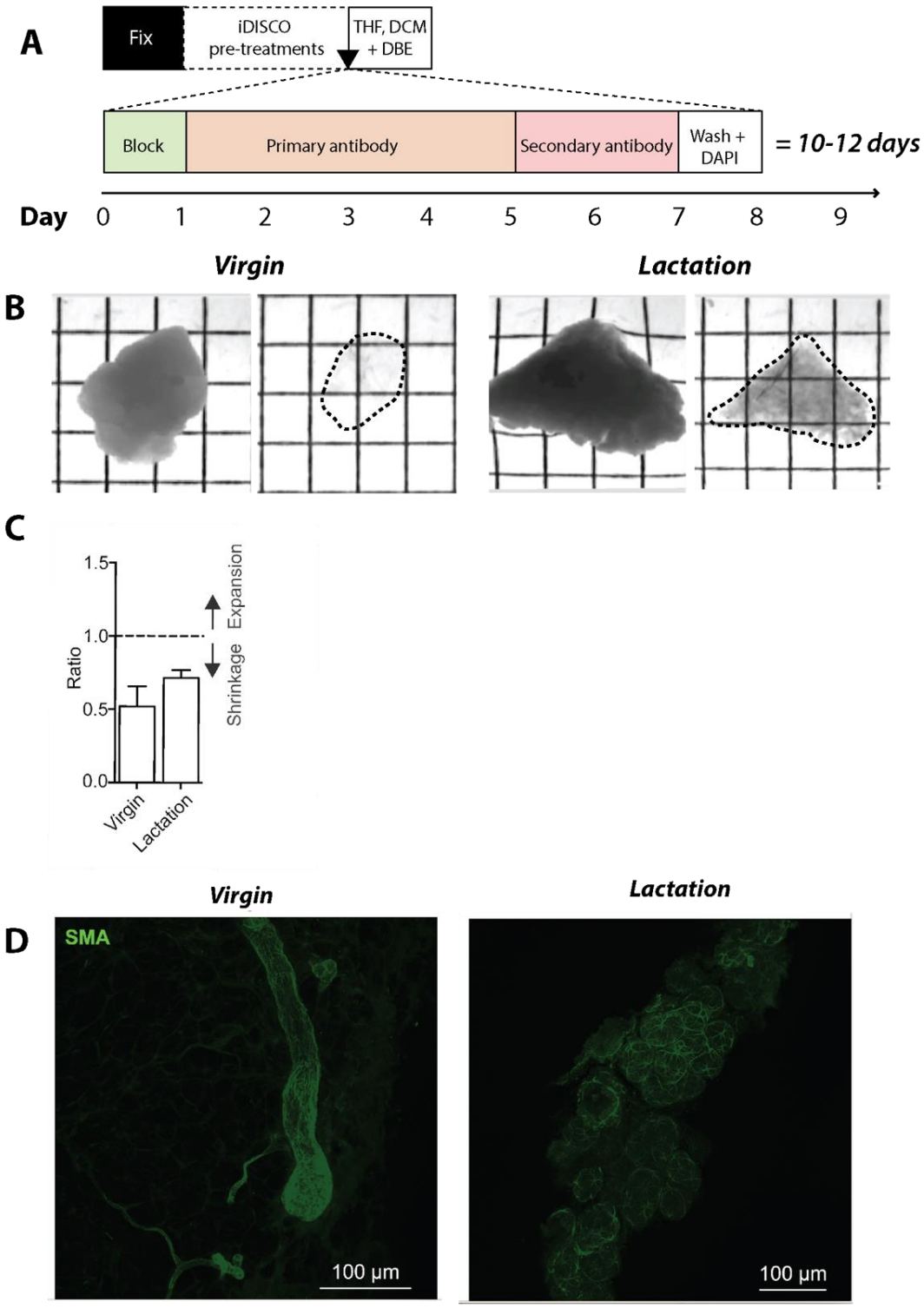


Figure 3.2.1.5 3DISCO-based optical tissue clearing in the mammary gland

(A) Schematic outline of 3DISCO optical clearing protocol, including primary and secondary antibody staining, resulting in a total protocol time of 10 days. **(B)** Use of 3DISCO in both virgin (left) and lactating (right) tissue resulted in excellent transparency. **(C)** Quantification of size changes after 3DISCO clearing in both virgin and lactating mammary gland tissue fragments (n=3 for each time point). **(D)** Fluorescent antibody-based staining with SMA was used in virgin mammary gland tissue. Tissues were imaged using confocal microscopy, and resulted in suboptimal and inconsistent staining at both time points.

3.2.2 Use of optical tissue clearing in cancer research

3.2.2.1 Optical clearing in mouse syngeneic tumour models

The performance of the clearing protocols was tested on *ex vivo* mammary gland tumours to determine whether the techniques described herein could be adopted for use in cancer research. TUBO tumours are derived from a cloned, HER2 overexpressing syngeneic tumour cell line originating from a spontaneous a BALB-neuT mouse mammary carcinoma (Rovero et al., 2000). Tumours were dissected out after reaching maximum growth limits permitted on the project licence, and then divided into small fragments (approximately $1cm^3$). We then compared staining of HER2, K8 and DAPI when using the three optical clearing protocols that worked best in the normal mammary gland - PACT-sRIMS, CUBIC and SeeDB. With all three protocols, we were able to image HER2-expressing cells at an enhanced depth and in high resolution (Figure 3.2.2.1a,b,c). K8 expression was also seen by fluorescence co-staining, but in a similar manner to that seen in the normal tissue, K8 was lower using the original CUBIC protocol compared to the other clearing methods (Figure 3.2.2.1b) - but this may be improved with the use of the second generation CUBIC R1a (Kubota et al., 2017).

We therefore opted to use SeeDB to examine the use of other antibodies in tumour imaging, due to its higher amenability to antibody staining (Figure 3.2.2.1c). In larger tumour fragments, we were able to see cells expressing the normal luminal marker K8, which appeared to be concentrated on the outer edges (Figure 3.2.2.2a), in addition to Her2 expressing tumour cells (Figure 3.2.2.2b). We also opted to co-stain for cleaved-caspase 3 (CC3), a marker of cell death, which has also been implicated in increased recurrence and death rates in breast cancer (Huang et al., 2011). This showed positive staining that was concentrated in the centre of lobules, as well as in between lobules (Figure 3.2.2.2a,b).

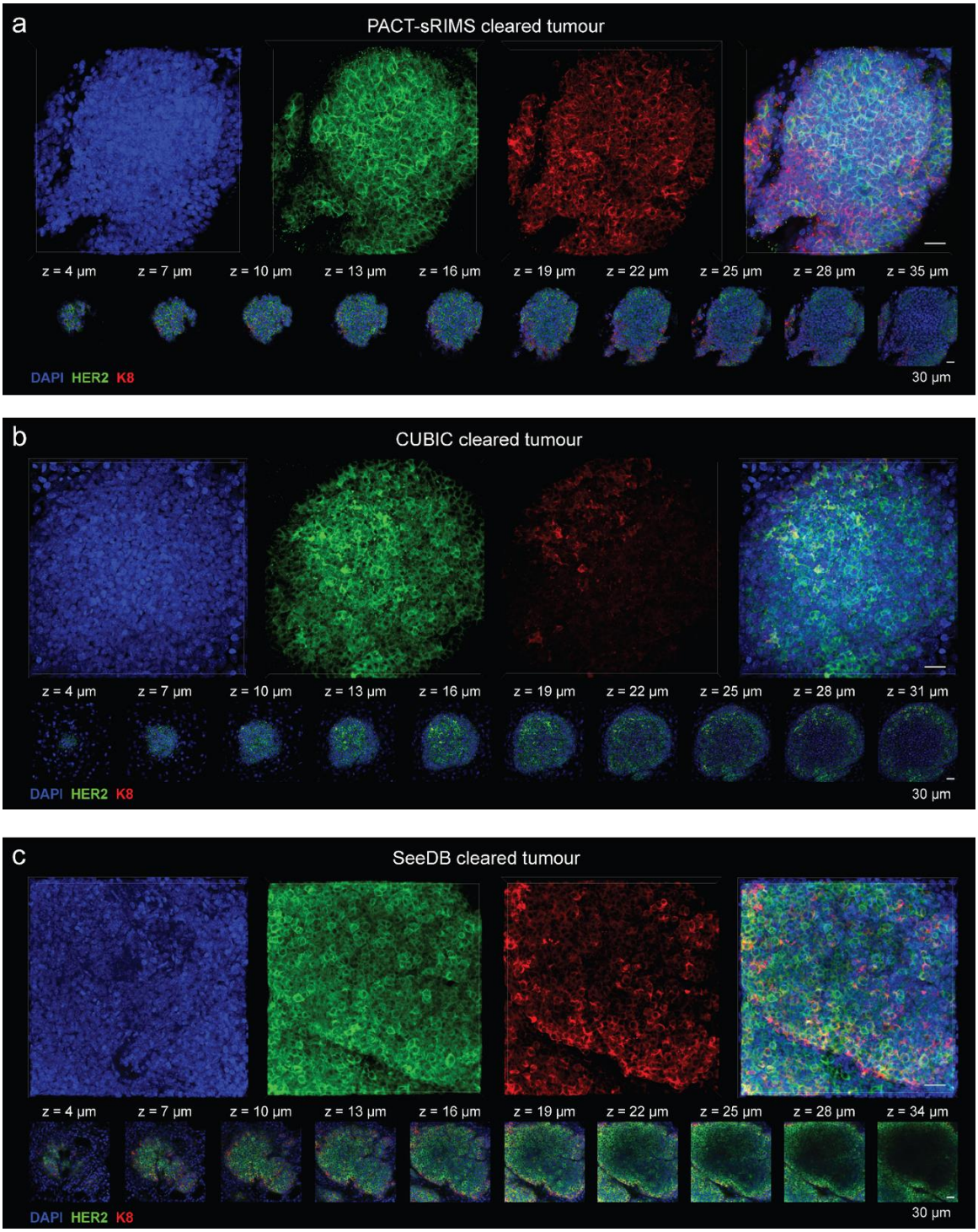


Figure 3.2.2.1 Assessment of optical tissue clearing protocols in mammary tumours

TUBO tumour fragments were cleared using PACT-sRIMS **(A)**, CUBIC **(B)** and SeeDB **(C)** optical clearing protocols. Tumours were stained for DAPI, HER2 and K8 and imaged using confocal microscopy. Images show maximum intensity projections with corresponding optical slices and their z-value for each protocol.

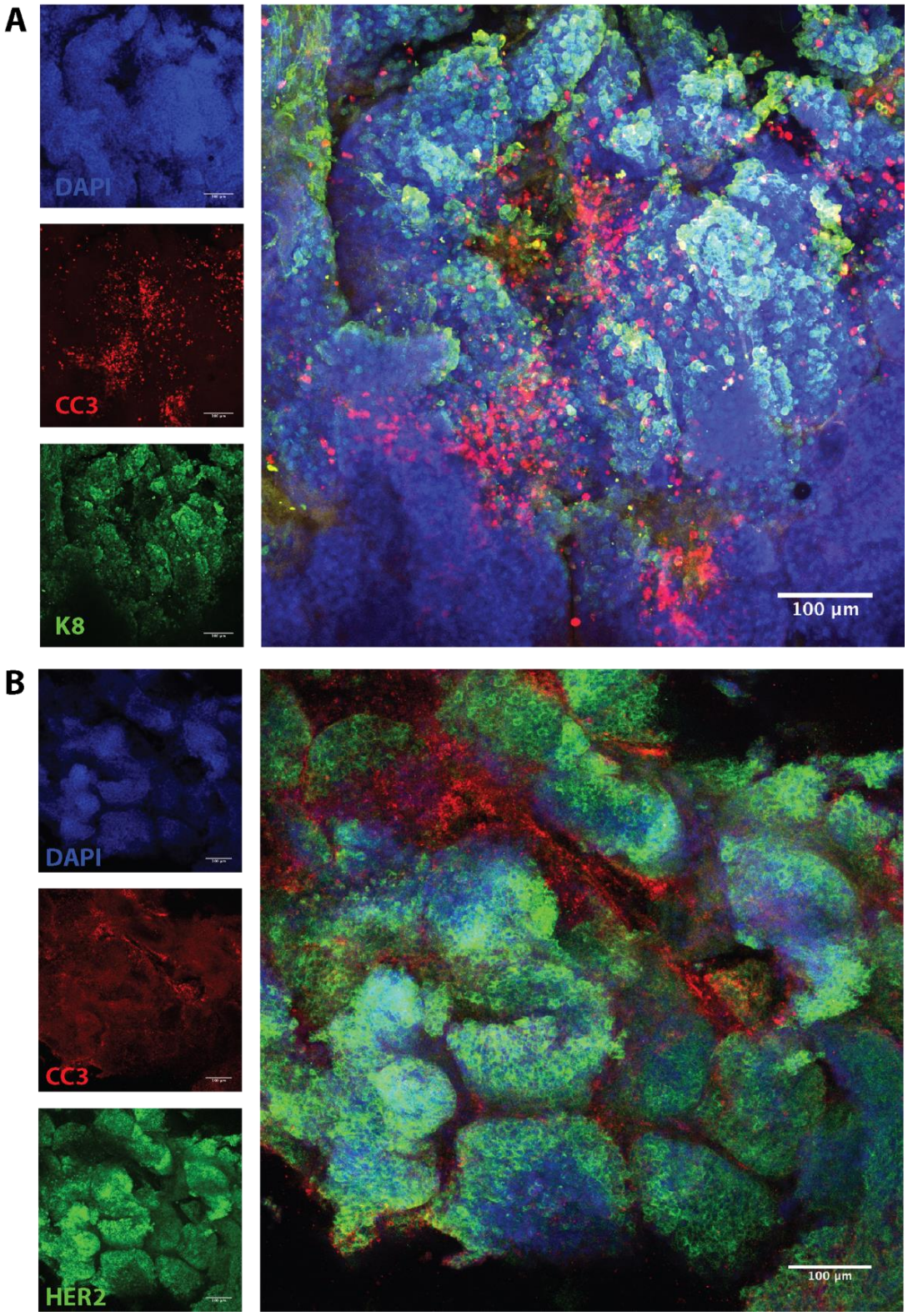


Figure 3.2.2.2 SeeDB optical clearing to examine apoptotic cell death in TUBO mammary tumours

TUBO tumour fragments were cleared using SeeDB optical clearing protocols and stained for DAPI, K8 and cleaved caspase 3 (CC3) **(A)** and DAPI, HER2 and CC3 **(B)** and imaged using confocal microscopy to examine apoptotic cell death. Images show maximum intensity projections with corresponding single colour images.

However, we observed reducing intensity of DAPI, HER2 and K8 staining with increasing imaging depth when looking through z-stacks of images. Initially, this was thought to be due to technical artefacts – i.e. suboptimal fixation or inadequate antibody penetration leading to edge effects, or reduction in signal intensity when imaging deeper. To investigate if this was a phenomenon of the wholemount staining and imaging, we opted to carry out traditional 2D immunohistochemistry on formalin-fixed paraffin-embedded (FFPE) tumour sections. Sections from both the edge (Figure 3.2.2.3a) and middle (Figure 3.2.2.3b) of the tumour mirrored the pattern that was seen in the wholemount images, showing reduced staining intensity of both K8 and HER2 in the centre of lobules. This similar pattern displayed in tumours stained and imaged in both 2D and 3D rules out inadequate antibody penetration or signal reduction due to wholemount imaging. However, inadequate fixation could still play a part.

As such, to explore this further and discern if there could be a biological reason for this, we next carried out CC3 staining in FFPE tumour sections to examine if this could be due to an increase in cell death in the centre of tumours due to necrosis. In doing so, we saw multiple zones of CC3 positivity, especially in the centre of tumour lobules within areas of low E-cad expression, in both sections from the edge (Figure 3.2.2.4a) and middle of the tumour (Figure 3.2.2.4b) - again, mirroring the pattern seen in wholemount SeeDB cleared tumours. Together, this suggests that the reduction in fluorescence intensity may also be, in part, attributable to the inherent biology of tumours, whereby centres of the tumour mass are often found to have high levels of cell death and necrosis, for example due to insufficient blood supply and the resulting hypoxic environment. This also highlighted the usefulness of carrying out 2D analyses in parallel to 3D imaging, especially in the case of characterising large, complex and/or heterogeneous tissues, such as tumours.

We next assessed the proliferation in the tumours by examining Ki67 staining – a marker of proliferative cells in the G1, S, G2 and mitosis phase of cell cycle, but not in quiescent or resting cells in G_0 (Gerdes et al., 1984). This marker is often used in prognostic settings in breast cancer using tissue sections (Inwald et al., 2013; Li et al., 2014), and as such the utility in wholemount tissues could be important. We observed high levels of proliferation throughout the tumour fragment, but especially on the leading edges and boundaries of the lobules (Figure 3.2.2.5, arrows). Co-staining with CC3 in the tumours showed a concentration

of apoptotic cell death in different areas to Ki67-marked proliferation – again, at the centre of the tumours (Figure 3.2.2.5, arrowheads).

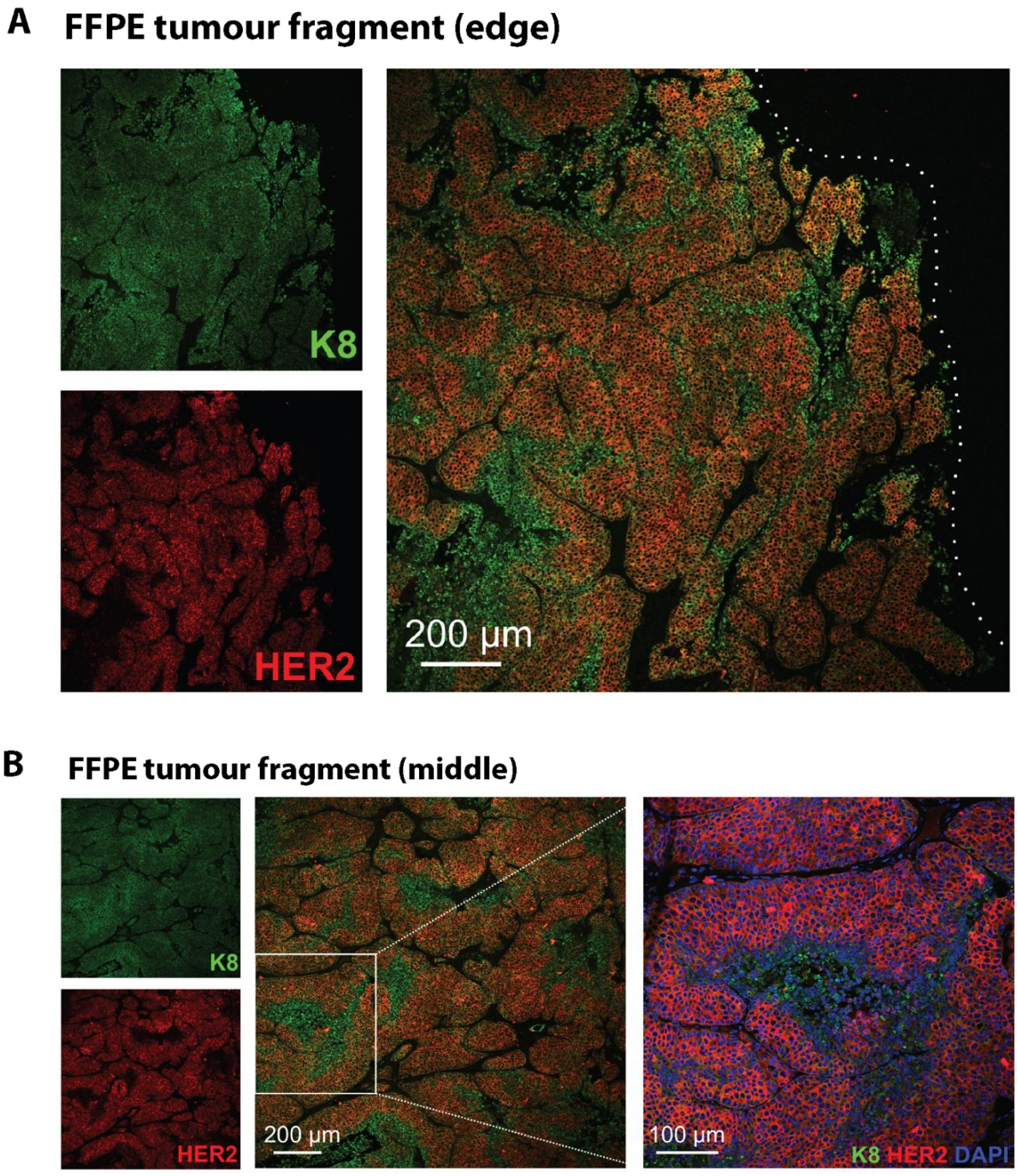


Figure 3.2.2.3 Staining pattern in FFPE TUBO tumour sections

TUBO fragments from the edge **(A)** and centre **(B)** of the tumour were formalin fixed and paraffin embedded, then sectioned and stained for K8 and HER2.

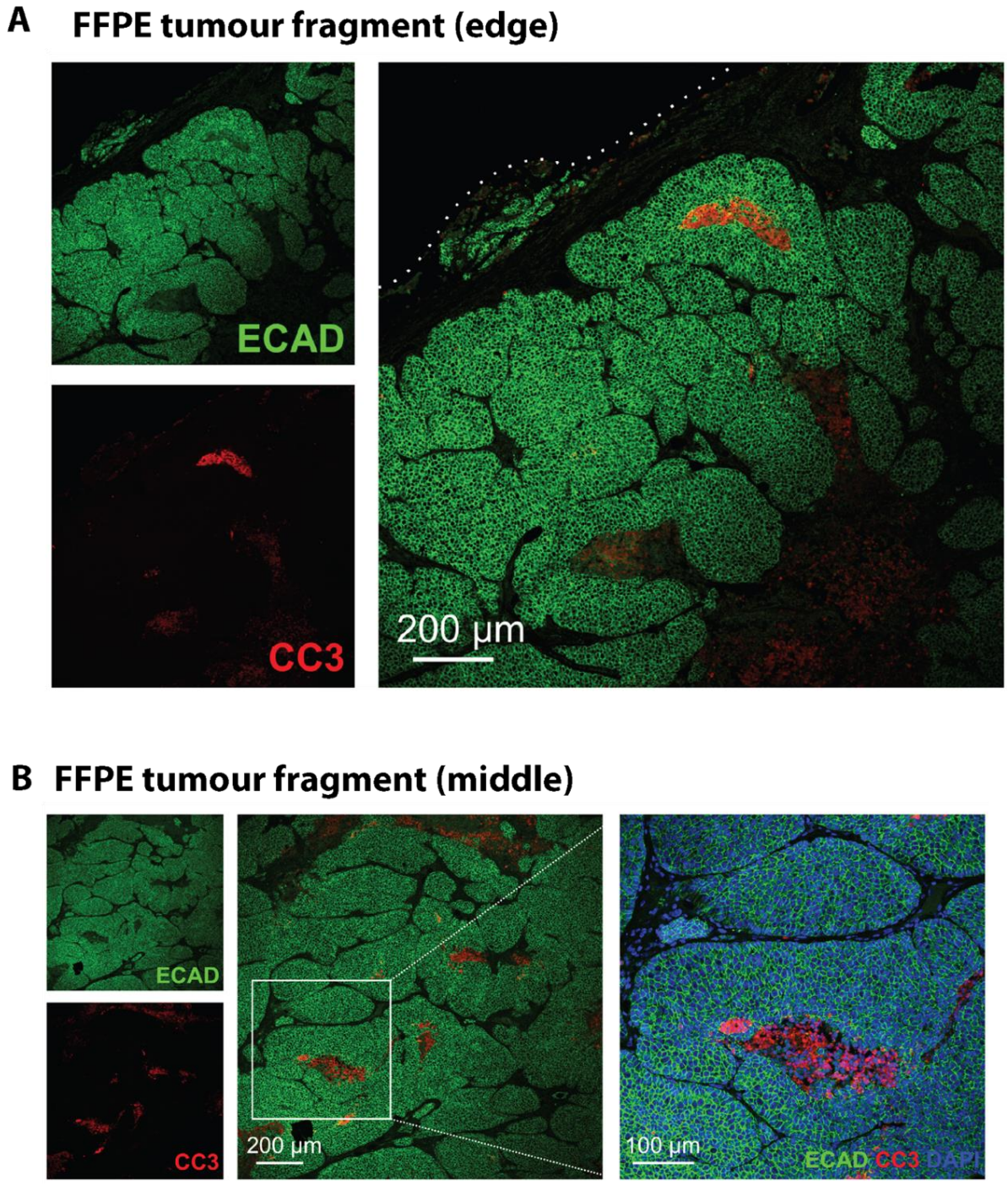


Figure 3.2.2.4 Apoptotic cell death in FFPE TUBO tumour sections

TUBO fragments from the edge **(A)** and centre **(B)** of the tumour were formalin fixed and paraffin embedded, then sectioned and stained for E-cad and CC3.

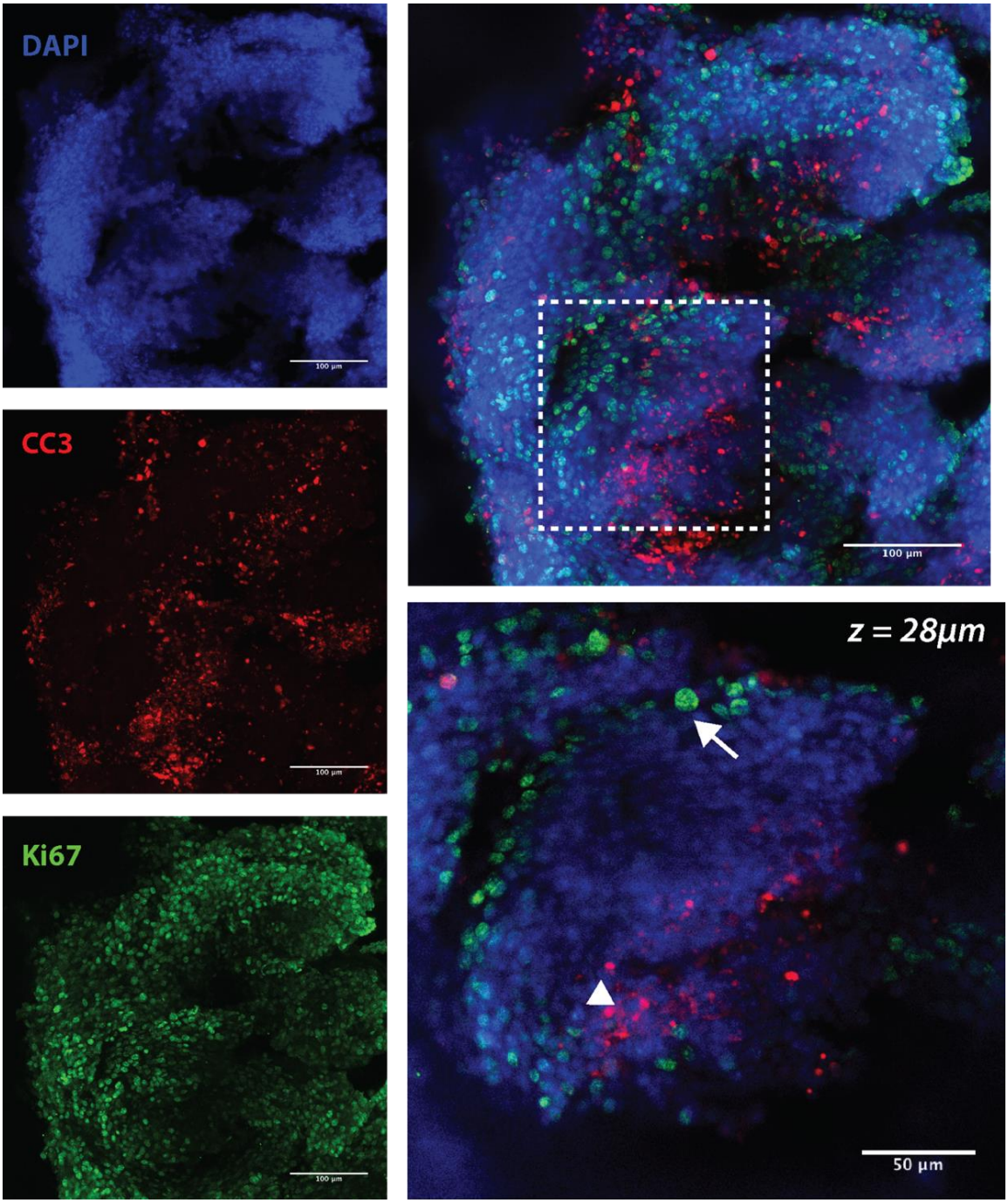


Figure 3.2.2.5 SeeDB optical clearing to examine cell proliferation and death in TUBO mammary tumours

TUBO tumour fragments were cleared with the SeeDB optical clearing protocol and stained for DAPI, Ki67 and CC3. Proliferation (marked by Ki67) appeared to be concentrated on leading edges of tumour lobules (arrows), whereas apoptosis (marked by CC3) was concentrated in the centre of tumour lobules (arrow heads). Images show maximum intensity projections, along with individual channels.

Since the TUBO tumour cells are a syngeneic line, and as such grow in an immunocompetent environment, we also wanted to investigate whether optical clearing could be utilised to examine the immune infiltrate in the tumours. This is a particularly important aspect of tumour research given the increasing understanding of the importance of the immune system in tumour development and maintenance, in addition to the number of immunotherapies in the clinic. Using 3D imaging in multiple focal planes, as opposed to 2D sections, allows a better appreciation of immune cell location in the context of an entire tumour. Additionally, as larger areas are imaged, it allows for patterns to be more easily seen, that may not be appreciated when surveying a single focal plane.

Therefore, we opted to stain for the markers CD3 and CD11b, which are pan-T and pan-myeloid cell markers, respectively (Reinherz et al., 1980; Arnaout et al., 1983). We were able to see zones of both CD3 and CD11b staining (Figure 3.2.2.6). Interestingly, these were often in between lobules of the tumour, and in different locations to one another. When examining at a higher resolution, CD3 cells with differing morphology were seen – with some exhibiting rounder (arrows) while others displaying more elongated shapes (arrowheads). This is potentially indicative of the different T cell subsets and activation status (Negulescu et al., 1996), however, co-staining with antibodies such as CD4 and CD8 would be required to confirm this.

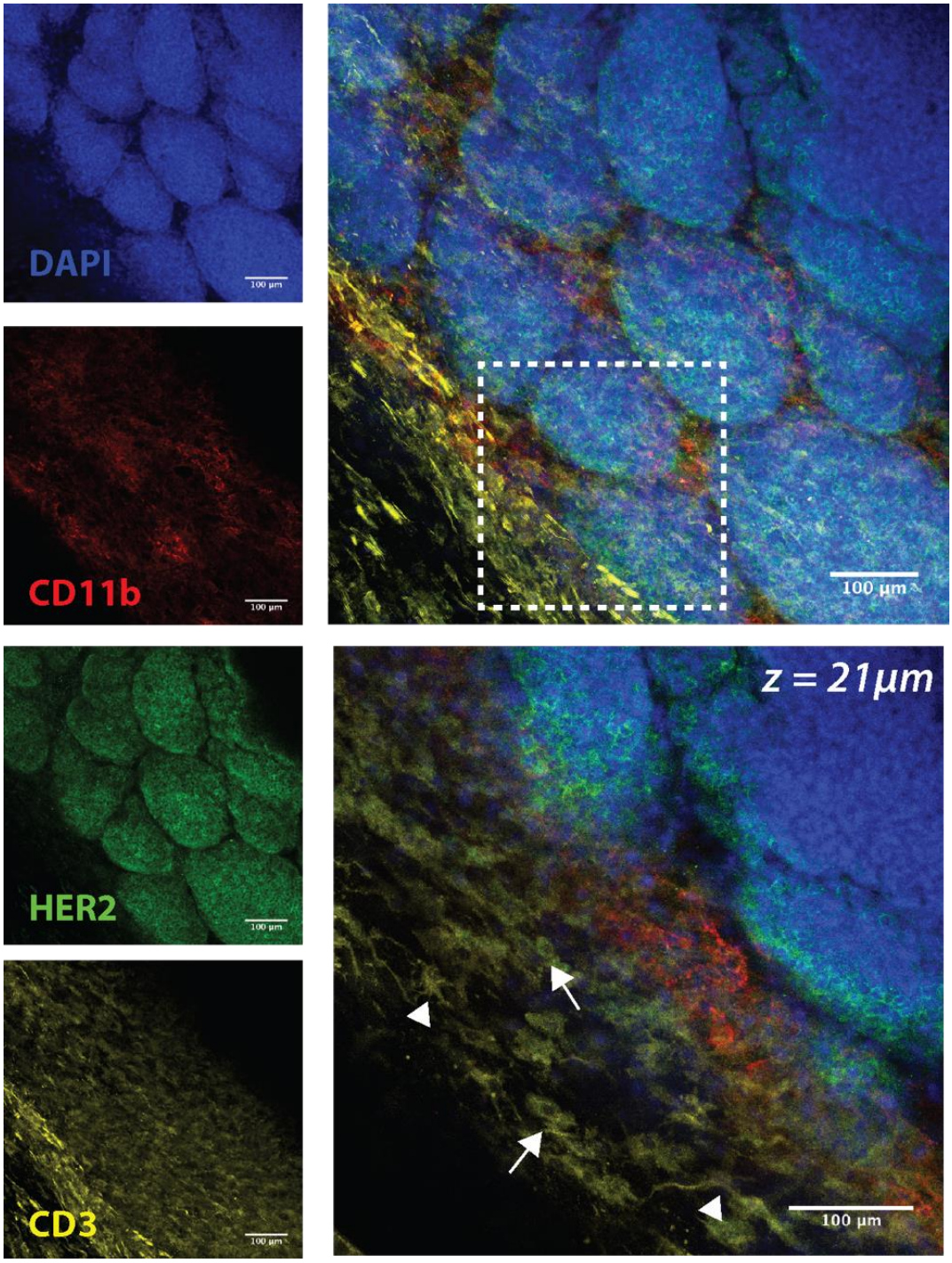


Figure 3.2.2.6 SeeDB optical clearing to examine immune infiltrate in TUBO mammary tumours

TUBO tumour fragments were cleared with the SeeDB optical clearing protocol and stained for DAPI, CD11b for myeloid cells, HER2 for tumour cells and CD3 for T cells. Images show maximum intensity projections, with individual channels, merge and magnified views (inset). Staining for CD3 reveals T cells with differing morphology, with some displaying an elongated shape (arrowhead) whereas others have a more rounded shape (arrows).

3.2.2.2 Optical clearing in human tumour samples and ex-vivo models

Whether CUBIC-based optical clearing could be used in human tumours – specifically, to investigate collagen involvement in carcinogenesis – was also tested (Hume et al., 2018a). In order to see if optical clearing could be used in the context of human tumours whilst also preserving the collagen architecture, a biopsy from an ER+ breast tumour was cleared using CUBIC R1a. This resulted in substantial optical transparency, with no noticeable differences in size (Figure 3.2.2.7a) (Hume et al., 2018a). The tumour fragment was then stained for K8 and ERα and imaged using 2-photon (2P) excitation microscopy, with the non-linear imaging method second harmonic generation (SHG) used to specifically detect collagen without the need for staining (Chen et al., 2012). In doing so, anisotropic collagen fibres could be observed running horizontally between ERα+/K8+ epithelial cells (Figure 3.2.2.7b), mirroring what has been seen in traditional 2D FFPE sections (Hume et al., 2018a).

An anisotropic collagen I-based 3D scaffold co-culture method had been previously developed in the laboratory for culturing both human breast cancer cell lines and *ex vivo* tumours (Campbell et al., 2017). This was expanded to more accurately recapitulate the human breast microenvironment, by using immortalised human mesenchymal stem cells (hMSC) isolated from breast reduction mammoplasties to provide an adipogenic environment (Hume et al., 2018a). Scaffolds seeded with the immortalised hMSC were differentiated into adipocytes and after fixation and immunostaining CUBIC optical clearing used. Samples were imaged again using 2P fluorescence microscopy with SHG for collagen I detection, in addition to coherent anti-stokes Raman spectroscopy (CARS), with the input signal adjusted for lipid detection (Le et al., 2009). This showed attachment of the hMSC's to collagen I pores, with a CARS signal on the perilipin+ vesicles (Figure 3.2.2.7c). Together, this shows that CUBIC-based optical clearing preserves collagen architecture in human tumour samples, as well as the relatively more delicate architecture of the scaffolds, in addition to maintaining the lipid vesicles (Hume et al., 2018a).

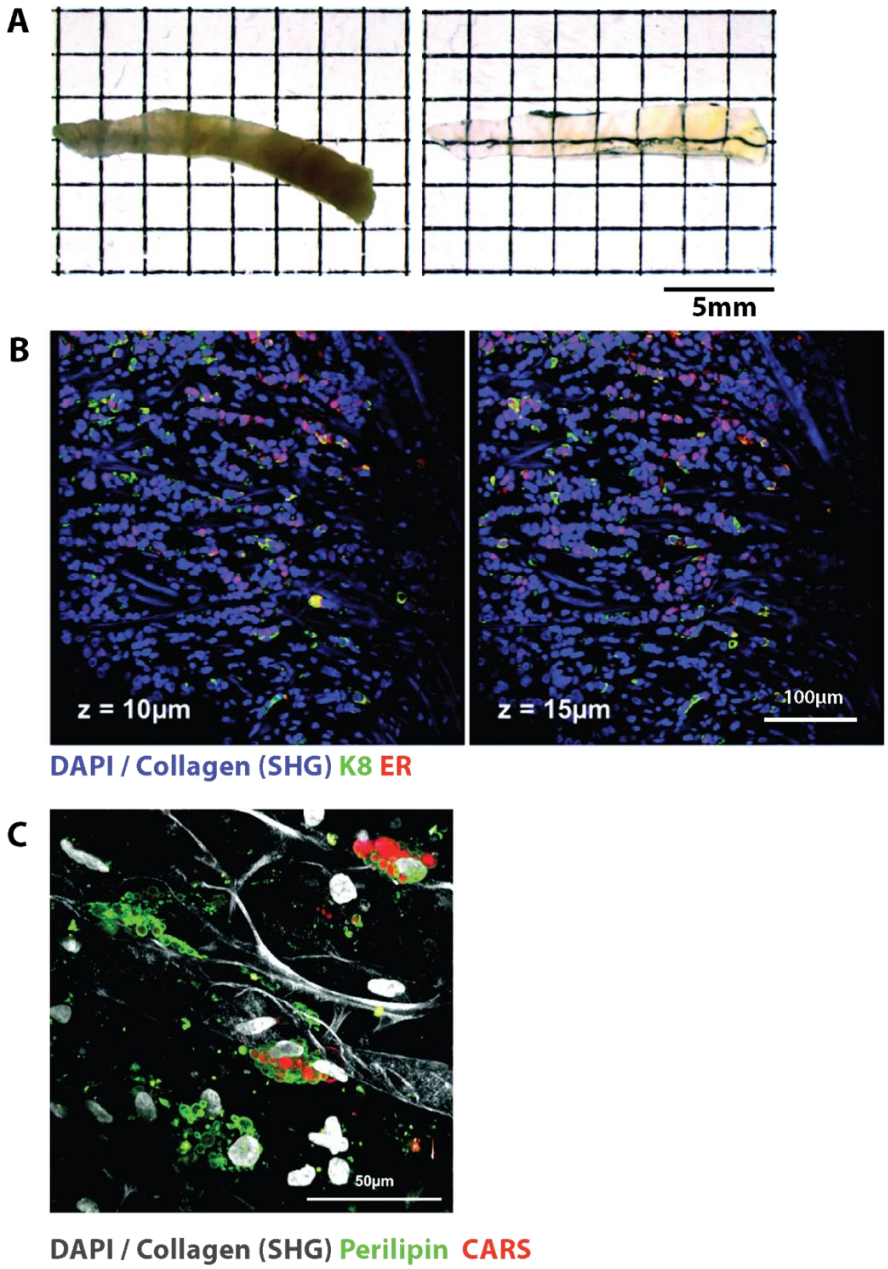


Figure 3.2.2.7 CUBIC optical clearing in human tumours and collagen co-culture models

(A) Use of CUBIC optical clearing in a human ER+ tumour biopsy resulted in excellent transparency. **(B)** Following CUBIC clearing, tissues were stained with fluorescent antibody-based staining using K8, ER and DAPI then imaged using 2-photon microscopy, combined with second harmonic generation to image collagen fibres. Images show optical sections throughout the tumour, with z values listed. **(C)** A collagen and immortalised human mesenchymal stem cell model for adipogenesis was tested. After using CUBIC clearing, scaffolds were stained for DAPI and perilipin, with second harmonic generation to image collagen and coherent anti-stokes Raman spectroscopy (CARS) to image lipid vesicles.

3.2.3 Combing optical clearing with other microscopy methods

Recent optical and computational advances have bought about an increase in microscopy options. Many of these are applicable to imaging of wholemount tissues due to the ability to optically section tissues. Although the use of confocal microscopy worked well with the aforementioned optically cleared samples and is often an accessible microscopy method, we wished to also investigate the utility of other microscopy techniques, and the compatibility of the clearing protocols was tested with these.

3.2.3.1 Two-photon excitation microscopy

The first alternative imaging method assessed was 2P excitation microscopy. This allows for deeper imaging penetration due to the longer wavelengths used in comparison to confocal – a useful utility when imaging wholemount tissue. This also results in other benefits such as less light scattering, leading to reduced out-of-focus photobleaching as well as faster imaging speeds. To asses 2P microscopy, we imaged a SeeDB-cleared tumour stained with Her2 and DAPI. We achieved imaging depths of approximately 1.2 × 1.2 × 0.1 mm of tissue in less than 15 minutes per individual channel (Figure 3.2.3.1). Both the speed and imaging depth was superior compared to confocal microscopy. However, when imaging multiple channels, 2P microscopy presented issues with spectral bleed through and as such resulted in less utility compared with the ease of imaging up to 4 channels on confocal, as standard.

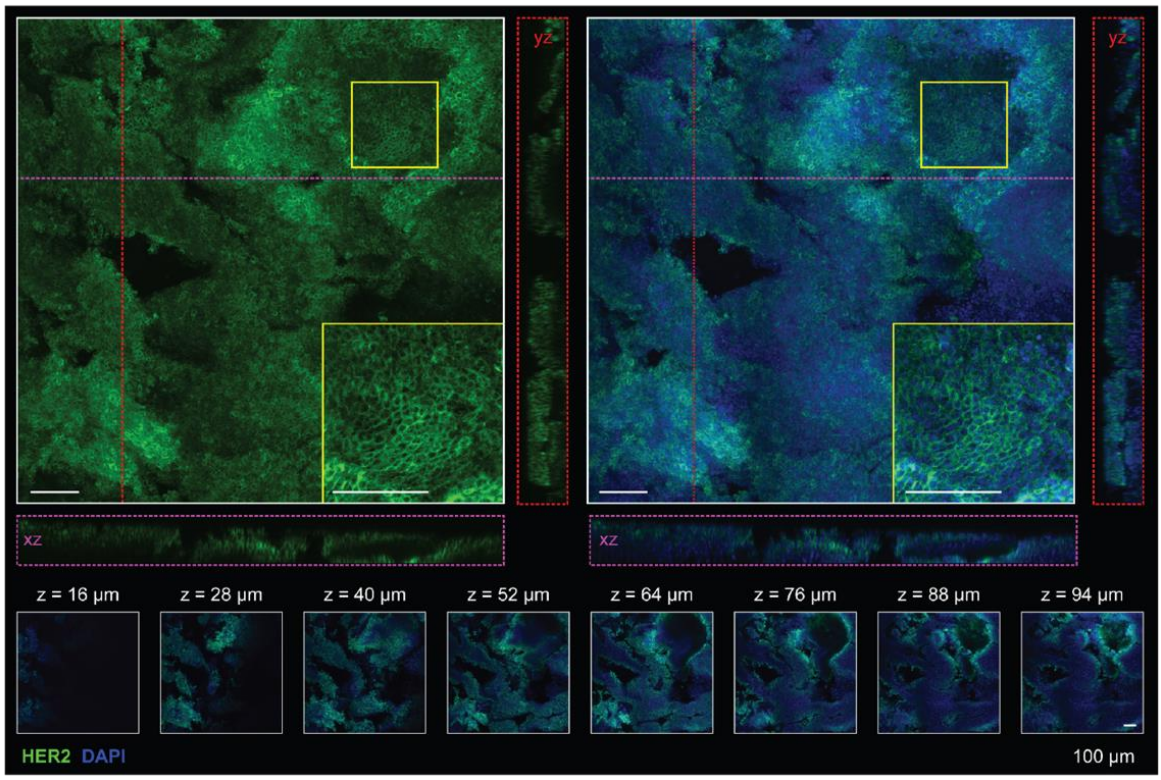


Figure 3.2.3.1 SeeDB optical clearing and 2-Photon excitation microscopy in murine TUBO tumours

TUBO tumours were immunostained with HER2 (green, left) DAPI nuclear staining (blue) and imaged using 2-Photon excitation microscopy. Example images show maximum intensity projections, with magnified views (inset) and optical slices (2 µm) and their corresponding depth (z value) relative to the first image in the image sequence (bottom panels). Orthogonal views, showing XZ (purple line and box) and YZ (red line and box) planes.

3.2.3.2 Light sheet fluorescence microscopy

The second alternative imaging technique assessed was light sheet fluorescence microscopy (LSFM). Here, a laser light sheet is used to illuminate only a thin slice of the sample, which reduces both photobleaching and background signal due to the increased optical sectioning capabilities. Additionally, due to the fact that samples are scanned using a plane of light (as opposed to a point, or pinhole, in confocal), image acquisition speeds can theoretically be up to 100-1000 times faster than point-scanning methods (Pitrone et al., 2013).

One important aspect of LSFM, where it differs from confocal, is the requirement for sample mounting. Unlike confocal, where the objectives are either above or below the sample, in the OpenSpim LSFM system used, the objectives and illumination comes from the side (Pitrone et al., 2013). As such, the samples require mounting in order to be suspended in the imaging chamber. One way to do this is to embed the tissue in agarose gel within a syringe, that can be suspended into the imaging chamber, surrounded by the objective media (schematic shown in Figure 3.2.3.2**)**. This therefore requires optimisation and adaptation for use with the optical clearing protocols to ensure that the matched refractive index is maintained between the sample, agarose, objective media and objective.

The first optical clearing protocol we tried with LSFM was SeeDB, due to its compatibility with fluorescent antibody staining. Tissue was put through the usual SeeDB clearing process and then suspended in a 1% w/v agarose solution in the syringe tube. However, the transparency of the tissue was lost due to the exchange of fructose into the agarose, and the water content returning to the tissue. Therefore, the agarose roll containing the tissue then required incubation with the fructose solutions again, which resulted in a lengthy process.

To combat this, we decided to combine the fructose solutions with the agarose to maintain a fructose concentration and limit the exchange of water and fructose. However, even at low fructose and agarose concentrations, the resulting gel became more solid and brittle. This lack of mailability presented technical issues when mounting in the small syringe system. Moreover, when imaging the tissue on the LSFM, the fructose/agarose gel appeared to reflect the light, thus preventing the tissue from being imaged. The reason for this could be due to the chemical structures of fructose and agarose monomers. When combined, these are able to form hydrogen bonds and produce a polymer, with a reduced elastic modulus (Nishinari et al., 1994; Maurer et al., 2012).

Therefore, we opted to try the combination with PACT-sRIMS, due to the fact that this also involves clearing with detergents, and as such does not only rely upon RI homogenisation to optically clear. Moreover, the sRIMS solution is very economical and can be used in the large quantities that are required for the imaging chamber. When combining agarose with sorbitol, the resulting gel both maintained the transparency of the tissue and was more amenable to tissue mounting than the stiff fructose/agarose gel.

We also combined CUBIC tissue clearing with LSFM. For this, tissue was cleared using the standard CUBIC protocol, staining with SMA and then mounted in agarose. Unlike SeeDB clearing, the tissue remained transparent, presumably due to CUBIC using detergents to optically clear and not relying only on RI matching. Using CUBIC R2 in the imaging chamber for a consistent refractive index allowed imaging of early gestation mammary gland tissue (Figure 3.2.3.3a) with a total depth of 800μm. Moreover, 0.8 x 0.8 x 1.5 mm of CUBIC-cleared normal virgin mammary tissue was imaged (Figure 3.2.3.3b), at acquisition times of less than 5 minutes per individual channel. Together, this highlights the importance of sample preparation in LSFM. However, although both imaging speeds and depths were superior to both confocal and 2P fluorescence microscopy, the time involved in sample preparation and optimisation meant that LSFM did not provide any benefits over confocal microscopy for routine imaging of the virgin mammary gland.

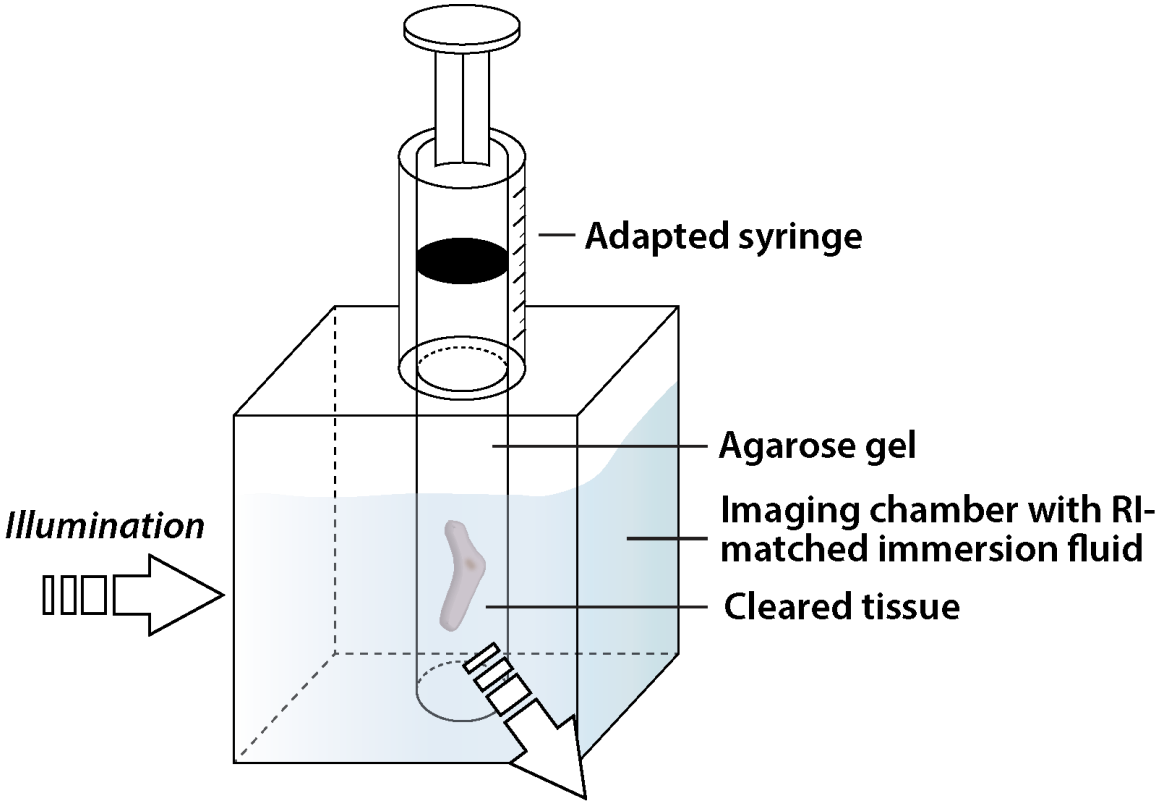


Figure 3.2.3.2 Schematic diagram of light sheet microscopy imaging set up

The optically cleared tissue is embedded in a low concentration agarose solution and placed in an adapted syringe. When ready to image, the agarose/tissue is then expelled into the imaging chamber which contains RI-matched solution (such as CUBIC R2). This is suspended to allow illumination, and can be turned to allow imaging from multiple angles.

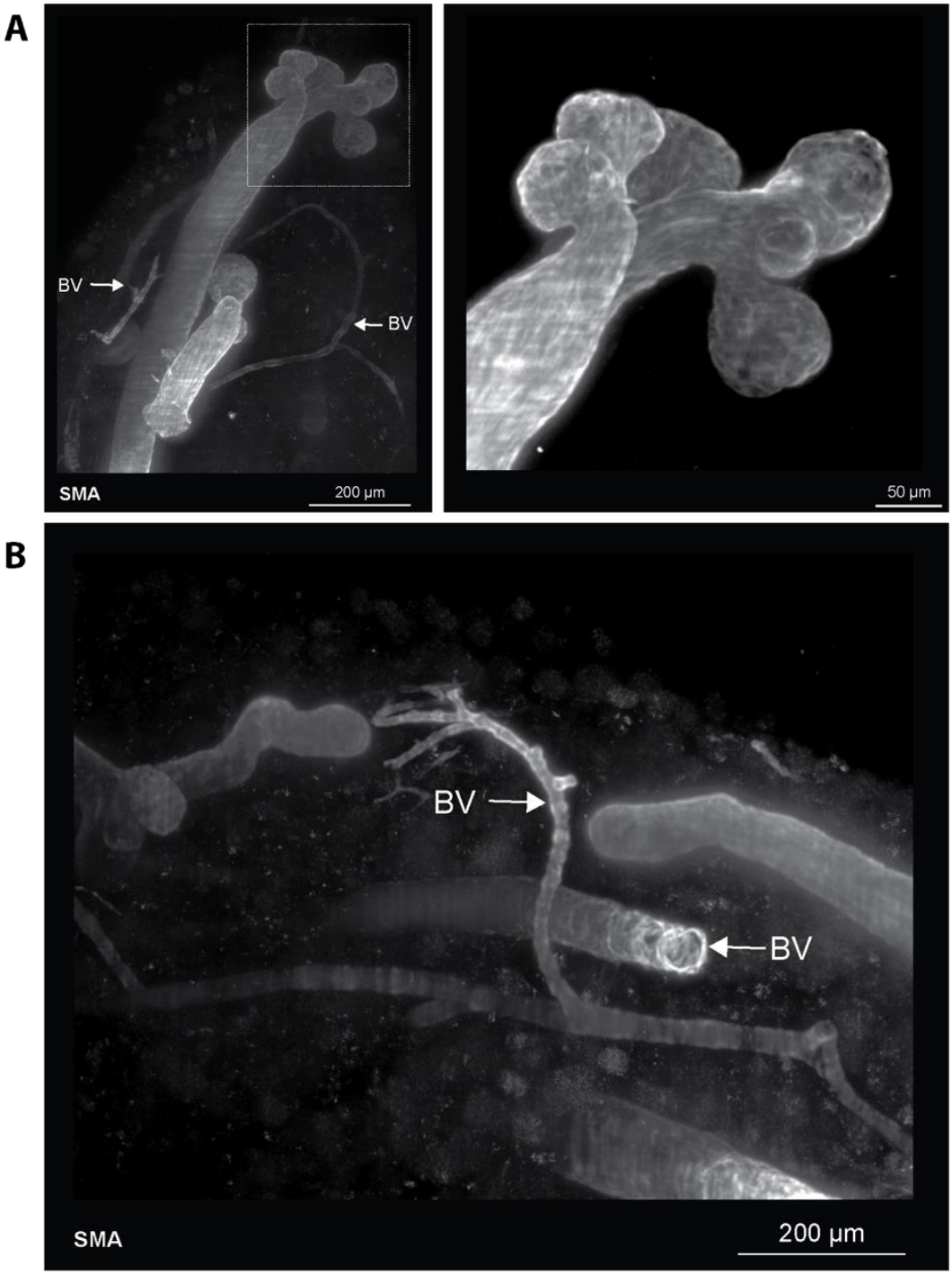


Figure 3.2.3.3 CUBIC optical clearing and light sheet fluorescence microscopy in the mammary gland

Early gestation (approximately day 5) **(A)** and virgin mammary gland **(B)** tissues were immunostained with SMA and cleared using CUBIC optical clearing. Using CUBIC R2 solution in the imaging chamber, tissues were imaged using light sheet fluorescence microscopy. Example images show maximum intensity projections of early gestation mammary gland **(A)**, and virgin tissue **(B),** with "BV" and arrows denoting SMA-expressing blood vessels.

3.2.4 Other applications of optical clearing methodologies

3.2.4.1 Liver

We decided to investigate the utility of optical tissue clearing in other systems. One tissue selected was the liver. Here, CUBIC, sRIMS and seeDB based clearing were compared. Due to the increased thickness of liver tissue compared to the mammary gland, liver tissue was sectioned using a vibratome into 400µm and 1000µm slices. Out of the protocols, CUBIC clearing was found to be the most beneficial, presumably due to the high haem content in the liver, and the pigment removal that CUBIC provides. The original CUBIC protocol was tested (Figure 3.2.4.1a)in addition to a recent addition, termed CUBIC L/R (Figure 3.2.4.1b) (see discussion). Both resulted in significant pigment removal and optical clarity, which allowed wholemount imaging using confocal microscopy with co-staining using antibodies for the structural protein osteopontin which marks ductal cells (Figure 3.2.4.1c)

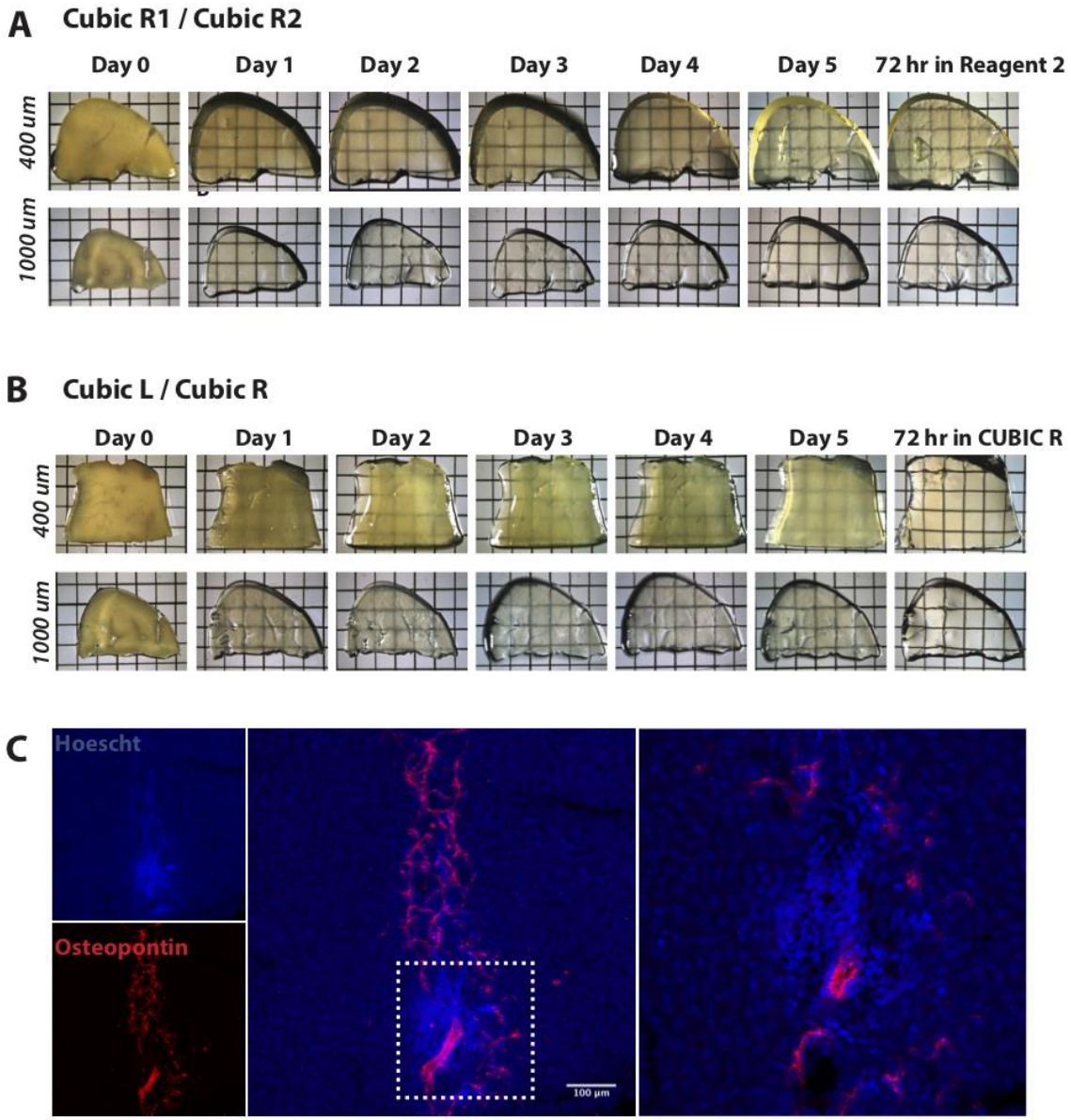


Figure 3.2.4.1 CUBIC-based optical clearing in murine liver tissue

Liver samples were sectioned to 400µm and 1000µm and cleared with either **(A)** CUBIC R1/2 or **(B)** CUBIC L/R. **(C)** CUBIC R1/2 samples were then stained for oestopontin and Hoechst and imaged using confocal microscopy. Images show representative maximum intensity z-projections with magnified region of interest.

3.3 Discussion

Altogether, the data presented in this chapter shows that optical tissue clearing can be successfully applied to study the mammary gland, both in homeostasis, during periods of development and lactation, in addition to in tumours of both the moue and human mammary gland. By assessing the leading protocols for use in the mammary gland, we have highlighted that not all optical clearing methods will work in every tissue. Moreover, these studies emphasise that some clearing protocols may be best suited to specific tissue types, with adaptation based on both the tissue in question and the microscopy technique used. For use in the mammary gland, this work concludes that SeeDB and CUBIC (specifically, the later published R1a protocol) are most suited for use in the mammary gland. An overview of the methods tested in the mammary gland is summarised in Table 1.

3.3.1 SeeDB optical clearing

Sucrose solutions have been previously used to reduce light scattering (Tsai et al., 2009). However, the authors of the original SeeDB protocol found fructose to be beneficial over other sugar solutions or water-based clearing agents – being both highly soluble in water, having a higher refractive index and causing less tissue shrinkage (Ke et al., 2013). One of the stated benefits of SeeDB is the preservation of fragile tissue architecture, such as dendrites in the brain. This also appears to be the case in the mammary gland, with preservation of fine features such as the striations of SMA in the basal cells remaining highly visible. Moreover, most antibodies tested worked with SeeDB clearing, suggesting that the protocol does not affect protein quality and epitope structure.

However, the downside to this gentle nature of SeeDB means that often the imaging depth is relatively limited compared to other protocols. This is especially true in the virgin gland, whereby the high proportion of adipocytes to epithelial components requires imaging through up to approximately 300μm of fat before reaching the ductal network. This is compounded by the fact that most commercially available and inbuilt imaging objectives have both short working distances, in addition to being built with the RI of air, water or oil in mind, which can result in spherical aberrations. However, the availability of specially manufactured objectives, as was developed and used in the original research protocol, which is matched to the RI of SeeDB, will allow for better compatibility with SeeDB optical clearing in the future.

3.3.2 CUBIC optical clearing

This work has shown that optical clearing using the CUBIC protocol is also highly compatible with the mammary gland. Despite being a reasonably simple, immersion-based protocol, we found excellent results in terms of both transparency and adaptability with microscopy techniques. However, one of the downsides found is the high level of background fluorescence seen, in addition to the reduced compatibility with some antibodies. Given that R2 contains 50% w/v sucrose, it may be that the Maillard reaction has a role to play in this, due to the interaction between the high level of sugar and proteins in the tissue. This could cause the formation of fluorescent by-products and reduction in protein quality (Dills, 1993) – both factors seen in CUBIC cleared tissue. As such, the inclusion of α-thioglycerol, as in used in SeeDB, may be beneficial in inhibiting the Maillard reaction and therefore reducing these subsequent issues.

We also used CUBIC and SeeDB for analysis of tumour infiltrate in 3D using optical clearing methods. This showed that these are viable methods for assessing tumours, which may be particularly important with the increase in immunotherapies for breast cancer. Recent optimisations and advancements of the CUBIC protocols may also provide even better optical tissue clearing solutions, for both normal tissues and other aspects of cancer research (Kubota et al., 2017; Tainaka et al., 2018). Using a similar chemical screen as was done in the original work, the authors have recently described the addition of CUBIC-L, which has been specially designed for de-lipidation. When combined with CUBIC-R, for refractive index matching, entire adult mice were able to be imaged to study cancer metastasis (Kubota et al., 2017) – providing further utility for optical clearing in cancer research.

Recently, further work has helped better understand the chemical properties required for successful optical tissue clearing. Consequently, an additional 5 CUBIC protocols have been presented, which are adaptable depending on required use. By combining 6 different CUBIC cocktails, users can develop a method based on the tissue used and required role – for example, de-lipidation, decolourisation, decalcification and refractive index matching (Tainaka et al., 2018). This highlights the possibility for combining different protocols and the importance of adapting methods based on the tissue type and research question.

3.3.3 Other clearing methods not tested

We also wanted to test an optical clearing protocol that utilises hydrogel embedding. Despite the relatively harsh treatment used in the PACT protocol, with high SDS concentrations used, most antibodies worked quite well. This could suggest that using hydrogel embedding is beneficial in stabilising tissue when applying optical clearing in the mammary gland. Although not tested here, it may be that applying a hydrogel embedding step prior to optical clearing methods which were not compatible, such as 3DISCO, would stabilise the tissue and result in a better outcome. Moreover, given that it is not feasible to carry out antigen retrieval methods on wholemount tissues, as is done on FFPE tissues, this makes it more susceptible to under- or over- fixation issues and subsequent reduction in protein and epitope quality. The addition of hydrogel embedding as the first step in other optical clearing protocols may be a useful addition to solving this issue.

The use of an organic solvent-based clearing protocol was also tested. Optical clearing with the 3DISCO protocol resulted in the excellent transparency of the tissue, and in less time in comparison with the other protocols. However, the quenching of endogenous fluorescence limited its future use in the mammary gland. Moreover, the significant shrinkage of the tissue and subsequent effects on the tissue integrity meant it was not compatible with the mammary gland.

However, a recent amendment of the 3DISCO protocol, "ultimate DISCO" (uDISCO), has been published, which utilises a different combination of solvents for better preservation of fluorescent signal (Pan et al., 2016). Moreover, uDISCO actually takes advantage of the tissue shrinkage caused by the 3DISCO protocol, allowing for imaging of larger volumes in a single scan - thus avoiding issues with sectioning and mosaic imaging. Use of uDISCO in an adult mouse brain resulted in a 55% reduction in tissue size, whilst preserving individual cell and vascular morphologies. Importantly, the new solvents included the addition of an antioxidant (α-tocopherol, or vitamin E), resulting in endogenous fluorescence being preserved over many weeks. If the use of uDISCO allowed for better preservation of tissue integrity than 3DISCO, the shrinkage of tissue and subsequent ability to image larger regions of the ductal network would be very beneficial. If the gland was reduced in size at the same proportions seen in the brain, it would also likely allow for imaging through the entire gland, from top to bottom.

Conversely, other recent protocols have utilised the increases in tissue size seen in optical clearing, resulting in the development of expansion microscopy (Chen et al., 2015; Murakami et al., 2018). Traditional optical microscopy is limited in its resolution to approximately 200 nanometers (nm), meaning that objects separated by less than half the wavelength of visible light cannot be separated (Simpson, 2006). As such, it is not possible to resolve anything below this diffraction barrier limit. In expansion microscopy, the expansion of the tissue is taken advantage of, by introducing a polymer network which can be subsequently swollen, increasing the tissue size with it (Chen et al., 2015). This allows for the capture of small nanostructures in the tissue that are not otherwise possible to image. This technique has enabled the imaging of objects at 70nm in both cultured cells and brain tissue, and even single molecules of RNA using fluorescent *in situ* hybridization (Chen et al., 2016).

In a similar manner, the CUBIC-X expansion microscopy and tissue clearing protocol utilised a combination of hyperhydration and subsequent gel embedding for the expansion of cleared adult mouse brain by 10-fold. This allowed subcellular imaging throughout the entire brain and lead to the development of the CUBIC-atlas – a mouse brain map with single cell annotation (Murakami et al., 2018). Although this was beyond the scope of the current work, the use of expansion microscopy in the mammary gland could be useful in providing new insights into the smaller details of the gland, such as investigating cell shapes and junctions. However, given the current limits brought about by imaging whole tissues, this would also require both more microscopy adaptation and data storage solutions to be developed.

3.3.4 Other microscopy methods not tested

The recent advances in microscopy have also bought about the opportunity to use optical clearing to enable imaging of whole organs. Here, we combined confocal, 2P and light sheet - fluorescence microscopy with optical clearing methods. However, all of these microscope set-ups have a certain level of adaptable configuration and as such, may result in different outcomes when tested in different settings or on different systems.

A key example of this is LSFM, with the open source system used, OpenSPIM, which allows for a high degree of adaptability depending on the required results (Pitrone et al., 2013). The set up tested here uses an immersion-based imaging chamber, and as such, some of the issues with compatibility with LSFM was due to the more complicated mounting methods required when compared to confocal. Moreover, the nature of the system also presents issues

regarding data storage due to the large file sizes that are the result of imaging such large tissues. There are a number of other microscopy systems that were outside the scope of this work, that would be compatible with optical clearing, including optical projection tomography.

3.3.5 Conclusions

This work has highlighted the compatibility of using optical tissue clearing for imaging the mammary gland, *in situ* and in 3D. These protocols can be utilised in future work in order to provide new insights into mammary gland development and were used for further work in this thesis.

Table 1 ***Comparison of optical tissue clearing methods in the mammary gland.***

Method	***Overview***	***RI***	***Clearing capability***	***Duration[a]***	***Preservation*** *structure*	*FP*	***Wholemount IHC***	***Storage***	***Rehydration and sectioning***	***Reference***
Uncleared	No clearing	–	–	1 day	Preserved	Preserved	Compatible	No	Possible	–
3DISCO	Organic solvent	1.56	Strong	2 days	Compromised	Rapid loss	Difficult[b]	No	Not possible	Ertürk *et al.* 2012
PACT-RC	Aqueous solution plus hydrogel embedding	1.45	Weak	10 days	Preserved, mild expansion	Preserved	Compatible	Yes	Not possible	Yang *et al.* 2014
PACT-sRIMS	Aqueous solution plus hydrogel embedding	1.46	Weak	13 days	Preserved	Preserved	Compatible	No	Possible[c]	Yang *et al.* 2014
CUBIC	Immersion, aqueous solution	1.49	Strong	5 days	Preserved, mild expansion	Some loss[d]	Semi-compatible[e]	No	Confirmed	Susaki *et al.* 2016
SeeDB	Immersion, aqueous solution	1.49	Moderate	5 days	Preserved, mild shrinkage	Preserved	Compatible	No[f]	Confirmed	Ke *et al.* 2013

[a]Duration from the time of tissue harvest (includes fixation time typically 6–16 hours for mammary tissue, but not antibody staining). [b]The fluorescence signal is rapidly quenched using benzyl alcohol benzyl benzoate (BABB) and a specialised imaging chamber is required for dibenzyl ether (DBE). [c]Likely to be compatible, but not tested in this study. [d]This is improved by using the second-generation protocol using Reagent 1A. [e]Some antibodies (i.e. E-cad, K8 and HER2) do not perform as well using the R1 protocol. This can be improved using the gentler R1a protocol. [f]Optimally imaged within 2 weeks but may be stored for several months. RI refractive index, FP fluorescent protein.

Chapter 4

Neutral lineage tracing in the mammary gland for studying development, homeostasis and stem cells.

4.1 Introduction

Unlike most organs, development of the mammary gland is a multistage process that occurs over several developmental stages, with the majority of development occurring postnatally. Once fully formed, it does not remain quiescent as both oestrous and pregnancy induce multiple cycles of expansion and regression throughout the reproductive life of female mammals. As such, the mammary gland represents a unique organ to study adult stem cells and their contribution to tissue development, maintenance and remodelling. The mouse is widely used to model mammary developmental mechanisms due to the functional similarities between the mammary tissue of both human and mouse, in addition to its amenability to complex genetic manipulations.

4.1.1 Mammary gland development

During mouse embryonic development, mammary placode formation is initiated at approximately embryonic day (E) 10.5, with the formation of two so-called milk lines from the overlying ectoderm (Propper et al., 2013). This is followed a day later with the asynchronous appearance of five pairs of placodes at specific and symmetric locations between the fore- and hindlimbs by E11.5 (Lloyd-Lewis et al., 2017). These develop asynchronously and invaginate to result in the formation of the mammary bud by E13.5, which evolves in to an epithelial tubular sprout by E15.5-16.5. The tubular sprout then invades the underlying mammary fat pad precursor which initiates the first phase of branching morphogenesis, culminating in the development of the primordial ductal tree by E18.5 (Howard, 2012).

From this point and during the weeks immediately after birth, growth of the ductal tree is isometric with body growth. However, the emergence of ovarian hormones during puberty then causes the second stage of ductal morphogenesis in the mammary gland. This ovarian hormone signalling combined with local growth factors in the microenvironment returns the gland to a highly dynamic state, whereby the rudimentary ductal structure expands to fill the fat pad, driven by cell proliferation in the TEBs (Watson and Khaled, 2008). These structures are present at the tip of ducts and comprise an outer layer of cap cells and a multi-layered inner core of body cells which form the two main cell lineages of the mammary epithelium, basal and luminal epithelial cells, respectively (Tiede and Kang, 2011). After branching morphogenesis is complete, the virgin mammary gland is thought to remain largely dormant,

with brief pulses of branching and alveolar budding and collapse in response to hormonal cues with each oestrous cycle.

The mammary gland undergoes another stage of branching morphogenesis, when it completely changes form again during pregnancy. Here, multiple signalling pathways are activated to orchestrate an unprecedented amount of proliferation during the process of alveologenesis. During lactation oxytocin is induced by suckling which stimulates the contraction of alveolar basal cells to expel milk from the lumen. When this ceases and the young are weaned, post-lactational regression of the mammary gland, known as involution, occurs, marked by extensive cell death of secretory alveoli (Kreuzaler et al., 2011), which returns the mammary gland to a structure akin to the nulliparous state (Watson, 2006).

4.1.2 Lineage tracing in the mammary gland

Use of lineage tracing to define stem cells is an oft-used technique with the advent of mouse genetic modification. Mouse models are engineered to express reporter proteins, often fluorescent or colourimetric, that are driven by pre-specified promoters for specific, pre-defined lineages. Once the reporter protein is induced in a cell, that cell and all of its daughters will harbour the genetic mark. If the cell marked is a stem or progenitor cell then this allows for the tracking and fate mapping of its progeny. The use of lineage tracing in many other tissues has allowed for a conclusive description and identification of adult stem and progenitor cell populations involved in tissue development, homeostasis, and even carcinogenesis.

In the mammary gland, the first inducible lineage tracing study challenged prevailing work on the contribution of multipotent MaSCs to postnatal gland development (Van Keymeulen et al., 2011). Using an inducible Cre recombinase driven by K14/K5 or K8/K18 promoters, basal and luminal cells and their progeny could be labelled, respectively. When inducing tracing postnatally, unipotent cells were traced throughout the different stages of development, suggesting that unipotent, rather than bipotent, stem cells drive development and homeostasis in the postnatal mammary gland (Van Keymeulen et al., 2011). Moreover, when the differentiation potential of both epithelial lineages was determined using transplantation assays it was demonstrated that label-positive basal cells, but not luminal cells, were capable of reconstituting a mammary gland, in line with previous findings that originally identified multipotent basal MaSCs. In turn, this strongly supports the notion that the experimental context of transplantation forces de-differentiation and reprogramming of basal MaSCs into

both epithelial lineages due to the non-homeostatic regenerative conditions. Conversely, in the intact gland *in vivo* only lineage-restricted unipotent stem cells appear to exist.

Controversially, more recent studies using similar methodologies have once again disputed the unipotent hypothesis and instead provided further evidence in favour of the original paradigm for the existence of bipotent stem cells in the adult mammary gland (Rios et al., 2014; Wang et al., 2015). Bipotent MaSCs were traced at a clonal level *in situ* and this study demonstrated that basal-labelled cells generated both cellular lineages, were long lived, and contributed not only to the major stages of morphogenesis in the postnatal gland but also in ductal tree maintenance during adult homeostasis(Rios et al., 2014). Since, another study has also provided evidence for the existence of unique, multipotent MaSCs that are marked by the expression of protein C receptor (Procr), a novel Wnt signalling target in the mammary gland (Wang et al., 2015). However, a key limitation of this approach is the reliance on promoter driven reporter expression, leading to assumptions regarding the molecular identity of MaSCs and progenitors, which therefore lends itself to bias and limitations in the cells that can be traced. Criticisms have also arisen when tamoxifen is used to induce promoter driven reporter expression (Asselin-Labat et al., 2010).

4.1.3 Role of mammary stem cells

Consequently, the remarkable capacity of the mammary gland for rapid growth and regeneration through multiple cycles of oestrus and pregnancy has been attributed to the existence of adult mammary stem cells (MaSCs). Although the identification and characterisation of these putative stem cells is a highly dynamic area of research that has extended over 60 years, there remains significant controversy and dispute regarding the number, molecular identity, anatomical location and hierarchy of these elusive adult MaSCs.

As a result, there are now two main schools of thought regarding the identity of MaSCs. The first, largely brought about through lineage tracing experiments, suggests that each of the main cell types is derived from their own unipotent, lineage-restricted stem cell (Van Keymeulen et al., 2011; Wuidart et al., 2016; Scheele et al., 2017). The second hypothesis suggests that bipotent MaSCs do exist, and are able to give rise to both luminal and basal cell lineages (Rios et al., 2014); as supported by the ability of single cells to repopulate both lineages in a cleared mammary gland. While both transplantation and lineage tracing approaches have suggested

potential MaSC populations, they have also revealed a number of conflicting results and conclusions.

4.1.4 Aims of this chapter

The main aims of this chapter are to characterise the requirements for adult MaSCs during normal tissue development of the mammary gland; during embryogenesis, at puberty and during maintenance of the mature gland, and finally during lactation. This information will be used to elucidate the identity, number, lineage potential and location of adult stem cells in the mammary gland. This will first be investigated during pubertal development, when MaSCs are functionally active, and this information will be used to inform subsequent investigations in the developed adult gland.

Due to the inherent biases of these two approaches discussed above, a neutral way to assess stem cell dynamics in a physiological context, that makes fewer prior assumptions and can be interpreted at various stages of development, is required to resolve the long running debate on the MaSC hierarchy. Consequently, we will combine two novel neutral lineage tracing models to unequivocally define the mammary stem and progenitor cells throughout the stages of mammary gland morphogenesis and homeostasis – embryogenic, pubertal, gestation and adult homeostasis.

Unlike many other adult organs, there are no definitive markers for MaSCs that can be utilised for lineage tracing. As such, and to avoid any bias, a combination of neutral approaches was employed to trace single cells using two mouse models. In the first, utilisation of a $[CA]_{30}$ cassette allows for stochastic labelling of single cells to reveal the active contribution of single stem/progenitor cells to the developing gland. This approach relies on the inherent instability of microsatellite repeats, resulting in "slippage" during replication to label rare cells. Accordingly, this leaves an indelible mark on the cell and, if this is a stem/progenitor cell, all of its progeny will also be marked. Importantly, genetic labelling in this model is exceedingly rare, thereby allowing the fate of a single-labelled cell to be traced with a high degree of confidence. This "slippage" event has previously been reported to occur at a rate of 1/17,000 cell divisions within the intestine and murine embryonic fibroblasts (Kozar et al., 2013).

However, the ability to time the initiation of tracing and to enable marking of quiescent cells necessitates the use of a second model. For this, R26R-Confetti mice will be employed (Livet

et al., 2007; Snippert et al., 2010). In a similarly neutral approach to the R26$^{[CA]30}$ model, the confetti model traces cells expressing a reporter from the ubiquitous Rosa26 promoter, with the benefit of a multicolour reporter that aids with the distinction of individual clones. The construct contains two tandem invertible DNA segments; upon administration of tamoxifen, Cre recombinase (also expressed from the Rosa26 promoter) is translocated to the nucleus and removes the loxP flanked 'STOP' roadblock. Subsequent stochastic inversion and excision recombination events create four random expression possibilities (nuclear green fluorescent protein (GFP), cytoplasmic yellow fluorescent protein (YFP), cytoplasmic red fluorescent protein (RFP), or membrane-bound cyan fluorescent protein (CFP)). Cells will thus be marked at random in one of four colours.

Proposed MaSC identities are often derived from FACS sorted cells, which removes important spatial information and has been shown to be affected by the dissociation process (Shehata et al., 2012). Alternatively, identity is sourced from lineage tracing data visualised using 2D sections and may not accurately reflect the true nature of the heterogeneous gland. To resolve this, optical tissue clearing methods developed in the previous chapter will be utilised. Thus, optical clearing permits the visualisation of mammary tissue to a level that has not been previously achievable, while retaining the fluorescence (or colourimetric) signal and important spatial information.

4.1.5 Importance

The mammary gland provides a unique opportunity to investigate epithelial development extra-embryonically that is not available in other tissues. Moreover, study of maintenance and turnover of this organ has important implications for other epithelial systems.

Elucidation of the normal MaSC hierarchy also has important implications for understanding the biology of the mammary gland in pathological conditions. The heterogeneity of breast cancer is complex and although there is now an increased understanding of molecular and pathological subtypes (Visvader, 2009; Pece et al., 2010), there is still much work required to get closer to understanding the cell(s) of origin of breast cancer. Given the proposed longevity and suggested ability of MaSCs to survive multiple waves of cell death in involution, they represent a logical candidate for a potential cell of origin of breast cancer (Howard and Veltmaat, 2013). Reponses to treatment are still highly unpredictable and the hope is that

better stratification of breast cancer can improve response to therapy and may lead to the development of more targeted therapies (Pfefferle et al., 2015).

4.2 Results

4.2.1 Tracing embryonic mammary stem cells

4.2.1.1 R26$^{[CA]30}$ SYNbglA

Previous work investigating development of the mammary gland has suggested that the stem cells involved in the formation of the embryonic mammary rudiment are multipotent (Van Keymeulen et al., 2011; Spike et al., 2012; Boras-Granic et al., 2014; Rodilla et al., 2015; Wang et al., 2015). This has been shown using both transplantation assays, in addition to more recent work using genetic lineage tracing models. However, despite the volume of work there is still much disparity regarding both the molecular identity and differentiation potential of embryonic mammary stem and progenitor cells (Wansbury et al., 2011). The first model we used to investigate embryonic MaSCs was the R26$^{[CA]30}$ beta-glucosidase (SYNbglA) model.

The R26$^{[CA]30}$ model exploits the inherent instability of microsatellite repeats to cause spontaneous, random frame shift mutations during DNA replication. For this, a $[CA]_{30}$ dinucleotide repeat tract is placed upstream of an out-of-frame reporter cassette (a modified, thermophillic SYNbglA) (McCutcheon et al., 2010; Kozar et al., 2013). During DNA replication, a mismatch repair (or "slippage") of the reading frame can randomly occur, potentially moving the reporter in-frame and thus triggering constitutive reporter expression in the "slipped cell" from the R26 locus (Figure 4.2.1.1a).

Accordingly, this leaves an indelible mark on the cell and, if this is a stem/progenitor cell, all of its progeny will also be marked. Importantly, genetic labelling in this model is exceedingly rare, thereby allowing the fate of a single-labelled cell to be traced with a high degree of confidence. If the slippage occurs during embryonic development, then this allows for tracing of stem or progenitor cells from this point onwards. The SYNbglA reporter, combined with CUBIC optical clearing and co-staining with methyl green (outlined in chapter 2), allows wholemount imaging and identification of clonal structures at a macroscopic scale, in relation to the entire gland.

Hence, we sought to investigate labelling events at 7 weeks of age, at the end of pubertal development of the mouse. The approximate timing of the slippage event can be inferred from the location of the labelled cells relative to the nipple, since the exact timing cannot be

determined in the $R26^{[CA]30}$ model. At the start of puberty, the ductal tree develops approximately 1mm proximal from the nipple region, expanding outwards into the fat pad. Therefore, using this predictable developmental process it was assumed that slippage had occurred in a stem or progenitor cell postnatally if clones were observed proximally from this point (Paine et al., 2016).

Using this model, a number of SYNbglA positive regions were seen, with variable clone size. Interestingly, labelled cells were often intermixed with unlabelled cells (Figure 4.2.1.1b), presumably from the progeny of an unlabelled stem or progenitor cell, suggesting the presence of multiple cells that contribute to ductal expansion. Some of these labelled regions contained large clones of cells, occupying multiple ducts in the developing gland - likely having arisen from a highly proliferative MaSC. In one particularly rare situation, uniform labelling occurred in all cells, which was seen in multiple tissues examined, including both the mammary gland and the intestine (Figure 4.2.1.1c), presumably due to slippage occurring in very early on in embryogenesis or even in a germ cell. Nonetheless, the majority of labelled regions seen were relatively small, presumably developing from a recent frame-shift mutation or from strand slippage in a restricted progenitor or differentiated cell.

However, in one case a large clone was observed encompassing more than 10 mm of the ductal tree, and appearing to begin at the nipple (Figure 4.2.1.2a, asterisk denoting nipple). Interestingly, this appeared to be comprised of labelled basal regions in close proximity to ducts with luminal clones. In order to more accurately identify the lineage of labelled cells we therefore carried out traditional 2D immunohistochemistry using DAB co-staining for lineage markers, which allows for accurate single cell resolution for lineage determination. This confirmed what was seen on the wholemounts, with both basal (Figure 4.2.1.2b) and luminal clones (Figure 4.2.1.2c) detected in the two regions. The expansive size of this clone combined with the starting location at the nipple region of the mammary gland may suggest that a bipotent MaSC was labelled at some point during embryogenesis. Although definitive confirmation is problematic, one hypothesis is that this bipotent embryonic MaSC could have then given rise to both a luminal and basal daughter cell that was deposited along the expanding ductal tree to later generate the lineage-restricted progeny during postnatal development.

On another occasion, we detected a fourth mammary gland that was comprised of a high proportion of SYNbglA labelled cells (Figure 4.2.1.3a). The vast density of labelling (Figure

4.2.1.3bi,ii,ii) meant lineage was difficult to distinguish and so again DAB lineage staining was carried out using 2D immunohistochemistry. Markedly, in contrast the other large clone the majority of labelled cells observed were of the luminal lineage (Figure 4.2.1.4). This suggests that even during very early embryogenesis of the mammary gland, MaSCs can display a degree of lineage bias and as such are not all bipotent.

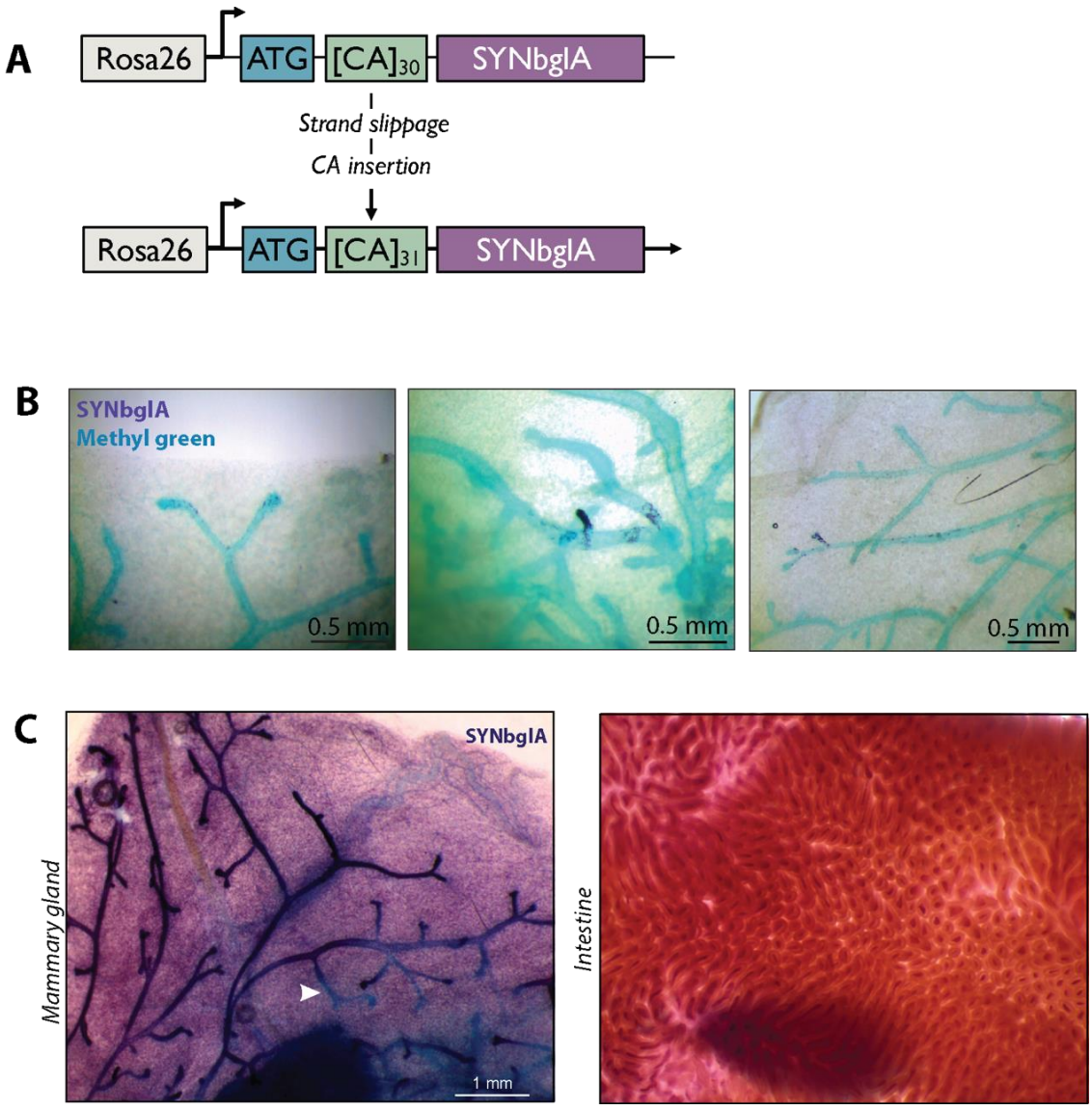


Figure 4.2.1.1 R26$^{[CA]30}$ SYNbglA tracing in the embryonic mammary gland

Diagram of the transcript of the R26$^{[CA]30}$ SYNbglA model. Random indels during DNA replication result in expression of the SYNbglA reporter, driven by the ubiquitous Rosa26 promoter. B) Multiple clones, marked by SYNbglA, can be seen when observing pubertal mammary glands. C) in one rare case, labelling of an embryonic germ cell resulted in uniform SYNbglA labelling in all cells, shown in the mammary gland and intestine (left and right, respectively)

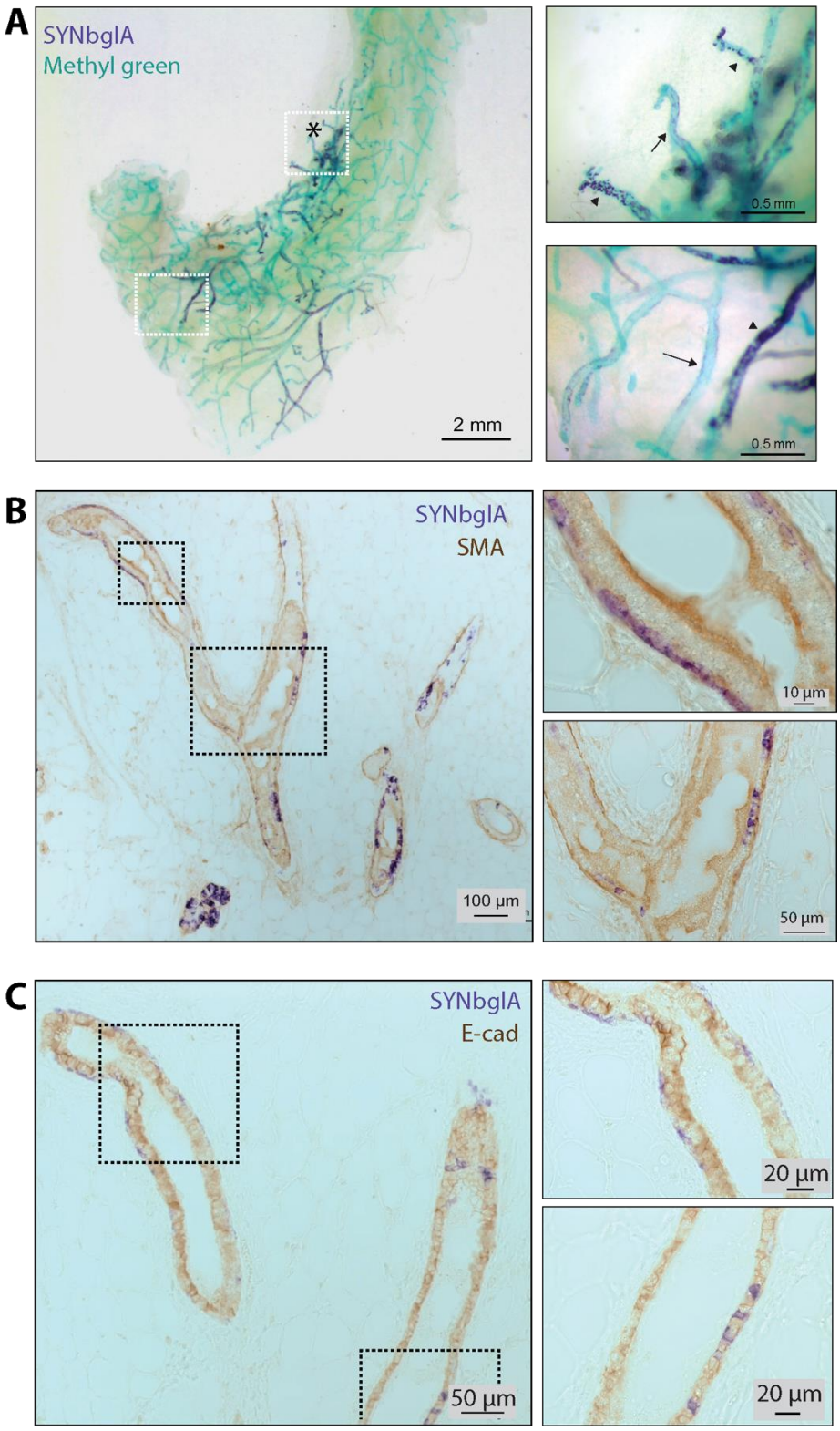


Figure 4.2.1.2 Clonal embryonic labelling in the R26$^{[CA]30}$ SYNbglA model

(A) A large clonal region visualised using stereomicroscopy, encompassing more than 10 mm of the ductal tree, labelled with SYNbglA and methyl green counter stain. Nipple denoted with asterisk. Labelled luminal and basal cells were observed in close proximity (arrowheads and arrows, respectively). Confirmation of lineage was carried out using immunohistochemistry with DAB staining for SMA, marking basal cells **(B)** and E-cad for luminal cells **(C).**

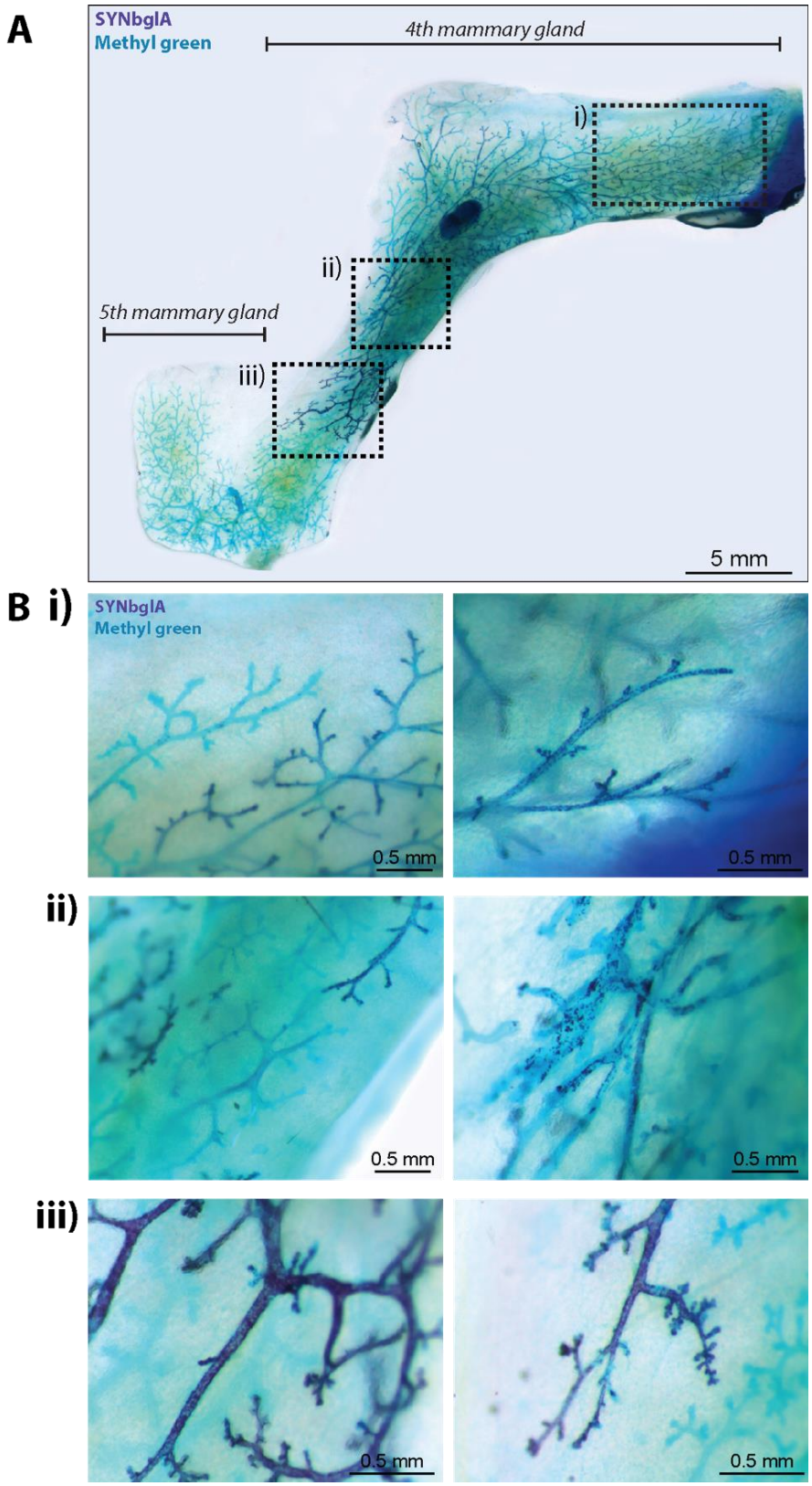


Figure 4.2.1.3. Clonal embryonic labelling in the R26$^{[CA]30}$ SYNbglA model

A) A large clone encompassing the majority of the fourth mammary gland, visualised using stereomicroscopy. B) Magnification of labelled regions shown throughout the gland, from the nipple region (ii) to the outer reaches of the fat pad (i and iii).

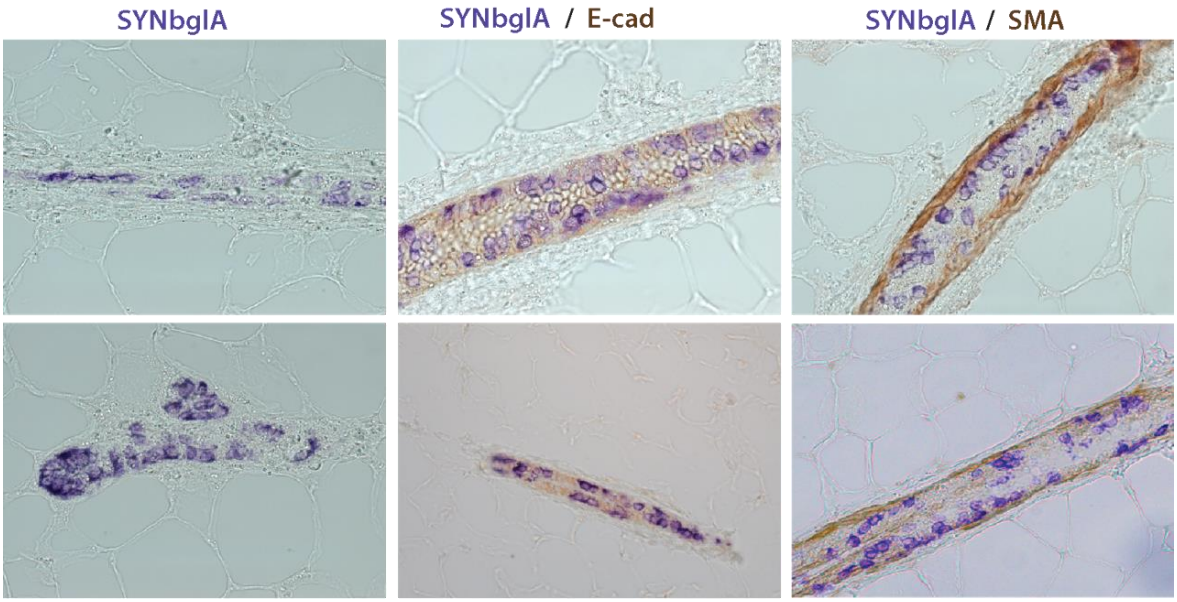


Figure 4.2.1.4 Immunohistochemistry of R26$^{[CA]30}$ SYNbglA clone shown in Figure 4.2.1.3.

Sections were co-stained for lineage analysis of the clone, with E-cad or SMA (for luminal and basal cells, respectively). Labelled cells (purple) all expressed K8.

4.2.1.2 R26-Confetti

The R26$^{[CA]30}$ SYNbglA model allowed for unbiased investigation of proliferative MaSCs that arise in embryogenesis and contribute to ductal development in puberty, without the need for tamoxifen. Furthermore, the extremely rare mutation rate provides assurance that clones are the progeny of a single stem or progenitor cell. However, the stochastic nature of the model also means that the exact timing at which tracing began cannot be accurately known. Moreover, it is not possible to precisely discern lineage *in situ* using the SYNbglA colourimetric stain, and consequently requires subsequent 2D immunohistochemistry.

We therefore sought to use a second unbiased approach for these reasons and to provide additional validation of the findings. Hence, the R26R-Confetti model was adopted. Similar to the R26$^{[CA]30}$ model, the R26R-Confetti model traces cells with reporters expressed under the control of the ubiquitous R26 promoter (Figure 4.2.1.5a). With administration of tamoxifen, Cre is translocated to the nucleus and removes the loxP flanked roadblock resulting in the random expression of one of four reporters (Snippert et al., 2010) (Figure 4.2.1.5b). This model can also be combined with the 3D imaging developed in the previous chapter, which is devoid of proteolytic digestion, and fluorescent immunohistochemistry which allows for *in situ* co-staining with lineage markers.

Accordingly, mice that were homozygous for R26Cre-ER^{T2} were bred with mice that were homozygous for R26R-Confetti (Figure 4.2.1.5a). The resulting progeny would therefore be hemizygous for R26CreER^{T2};R26R-Confetti, allowing lineage tracing following the administration of tamoxifen. Due to tamoxifen-mediated effects on oestrogen signalling, which can affect successful labour and result in pup mortality (Lizen et al., 2015), tamoxifen was delivered via oral gavage. Although tamoxifen is then is subject to the first-pass effect, resulting in reduced systemic bioavailability due to absorption by the liver (Rowland, 1972), this route of administration is reported to result in more uniform Cre-recombination due to systemic diffusion rather than a bolus dose given with intraperitoneal (I.P) injections (Park et al., 2008), in addition to reduced pup fatalities (Lizen et al., 2015). As such, a single low-dose of tamoxifen (equivalent to 33µg per g of maternal body weight) was administered by oral gavage to pregnant mice on approximately E16.5-E17.5 to induce neutral labelling of the embryos *in utero* (Figure 4.2.1.5c).

To investigate the involvement of the progeny of stem or progenitor cells labelled during embryonic mammary development in pubertal mammary morphogenesis, labelling events in the mammary gland were investigated at 6 weeks post birth. In order to analyse the entire focal plane without enzymatic tissue digestion, CUBIC optical tissue clearing (outlined in chapter 2) was used to image wholemount glands using fluorescent immunohistochemistry and confocal microscopy. Glands were stained for the basal cell marker SMA and confocal microscopy used to image endogenous fluorescence (YFP, GFP, RFP and CFP). Importantly, ductal morphology was similar when compared to non-injected animals not exposed to low-dose tamoxifen *in utero* (as seen in wholemounts in Figure 4.2.1.6).

Using this approach, we were able to observe large regions of labelled cells, with some spanning from the nipple to the outer reaches of the fat pad (Figure 4.2.1.6,a-c). So as to investigate the labelling efficiency, we analysed the patterns of labelling events in regions either proximal or distal to the nipple. Labelled regions proximal to the nipple were always comprised of at least two colours (Figure 4.2.1.7.a). However, distal regions were more commonly comprised of single colours (Figure 4.2.1.8a,b). To quantify this, the number of single- and multicoloured branches in the regions proximal or distal to the nipple were counted. 100% of nipple regions observed were multicolour (Figure 4.2.1.8c,i), whereas only approximately 40% of distal branches contained more than one colour (Figure 4.2.1.8c,ii). This supports the notion that as the gland develops, progeny of MaSCs become diluted out in distal branches. However, this also highlights the potential of clone convergence, whereby the progeny from separate stem / progenitor cells are labelled in the same colour and thus are unable to be distinguished as separate clonal populations.

Next, the lineage of labelled cells was investigated by observing labelled cell morphology and co-localisation with the basal cell marker, SMA. Mixed lineage, same-colour clones were occasionally observed in regions proximal to the nipple (Figure 4.2.1.7a, arrows), and to a lesser extent in distal regions (Figure 4.2.1.9a, arrows). These data support the previous hypothesis that embryonic MaSCs are multipotent and can contribute to both basal and luminal lineages. However, despite the high risk of clone convergence, single lineage clones were often observed in both distal and proximal regions of the ductal network.

A nearest-neighbour analysis was therefore carried out to quantify this, whereby the lineage of the nearest cell was counted. Again, the location of labelling events was analysed in relation

to the nipple (proximal or distal) in order to determine when labelling occurred. During expansion of the mammary gland during puberty, ductal elongation and side branching occur due to stem/progenitor cell proliferation within TEBs. As TEBs elongate, the progeny of these cells are dropped along the length of the developing ducts, with mixing of clonal progeny. Consequently, clones arising from labelling in both the embryonic and pubertal epithelium are thus not contiguous. Previously, proliferation was thought to exclusively occur in the TEBs (Wang et al., 2015; Scheele et al., 2017). However, work discussed in Chapter 5 of this thesis indicates that significant proliferation can also occur in the ductal region during pubertal development. Therefore, a method of clone analysis was carried out that did not exclude the possibility of ductal proliferation and took into account the *bona fide* 3D nature of the mammary epithelium.

Briefly, z-projections of randomly selected 3D image stacks containing label-positive cells were created. For all cells within each region, the lineage of the closest same-colour neighbour was recorded as either 'same' or 'different' by manual scoring. GFP, YFP, RFP and far red channels were imaged for each image sequence. However, it is also important to consider that using this method, two luminal labelled cells counted as "same" may in fact be distinct labelling events derived from different precursor cells; as such, this model may over-represent the percentage of possible unipotent clones. Similarly, a luminal and basal labelled clone counted as "different" may also be from distinct labelling events and thus this model may also overrepresent the proportion of possible bipotent clones; however, this was extremely small in this study.

This analysis showed that the majority of same-colour neighbours were in fact cells of the same lineage in both nipple regions and distal branches (Figure 4.2.1.9b,i and ii, respectively). This expands the finding seen with the $R26^{[CA]30}$ SYNbglA model - namely that embryonic MaSCs may not all be bipotent at E16.5-17.5. Indeed, this is supported by recent work that suggests that even as early as E14.5, embryonic MaSCs can display lineage restriction (Lilja et al., 2018; Wuidart et al., 2018).

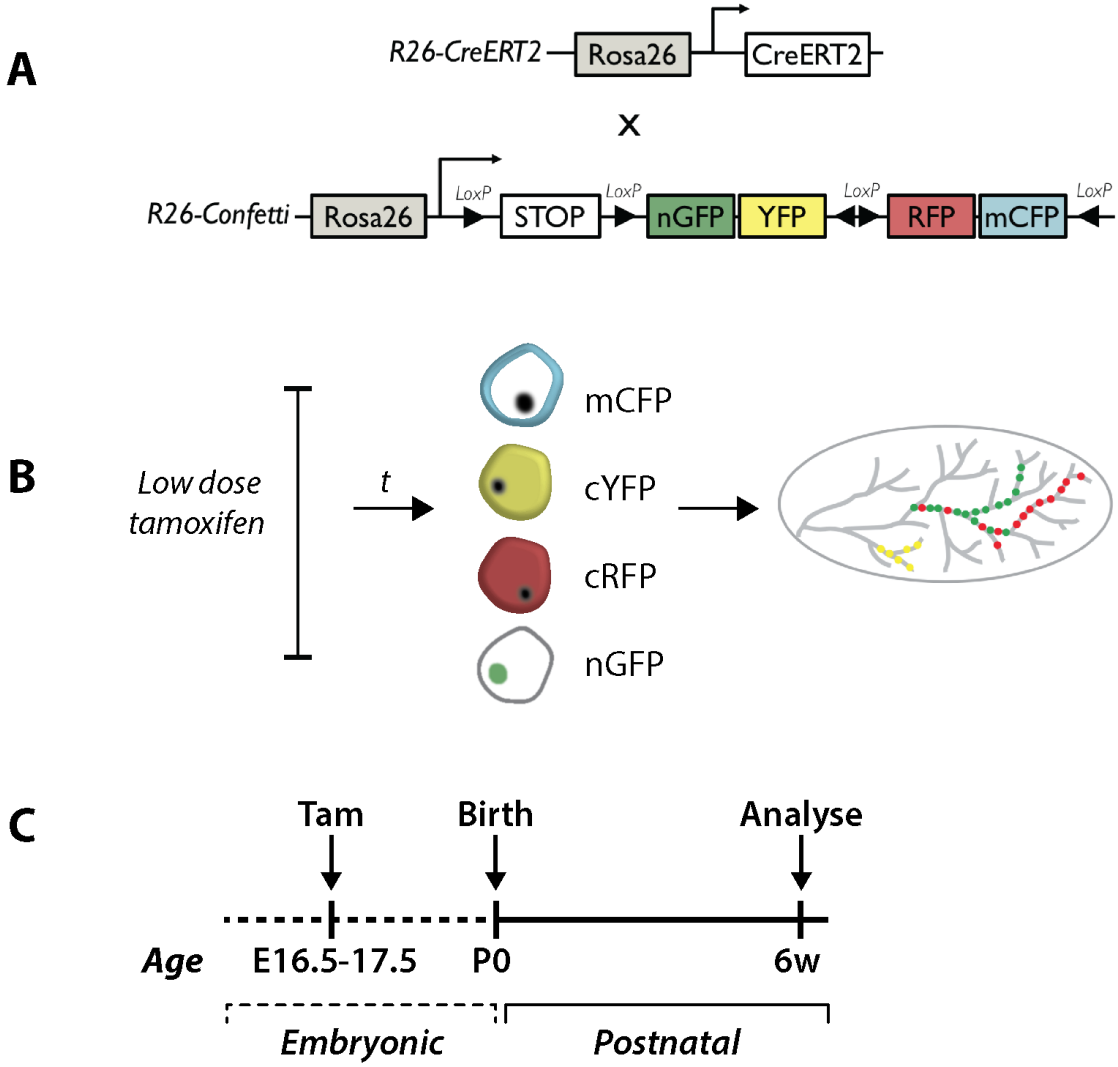


Figure 4.2.1.5 Overview of the R26-Confetti model for embryonic tracing of mammary stem cells

(A) Schematic outline of the construct for the R26CreER^{T2};R26-Confetti model. R26CreERT2 mice (expressing inducible Cre-recombinase in all cells) were crossed to R26-Confetti mice (expressing a conditional multicolour reporter in all cells) to generate double hemizygous mice. **(B)** Administration of low-dose tamoxifen produces stochastic genetic labelling of cells at low density. Labelling outcomes include membranous CFP, nuclear GFP, cytosolic YFP or cytosolic RFP. **(C)** Outline of lineage tracing using the R26-Confetti model for embryonic lineage tracing. Tamoxifen (equivalent to 33 µg/g of maternal weight) was administered by oral gavage on day E16.5-E17.5 (approximately 4-5 days prior to birth - mice on C57Bl/6 background with total gestation time of approximately 21 days). Tissue was harvested from R26CreERT2;R26-Confetti offspring at 6 weeks of age.

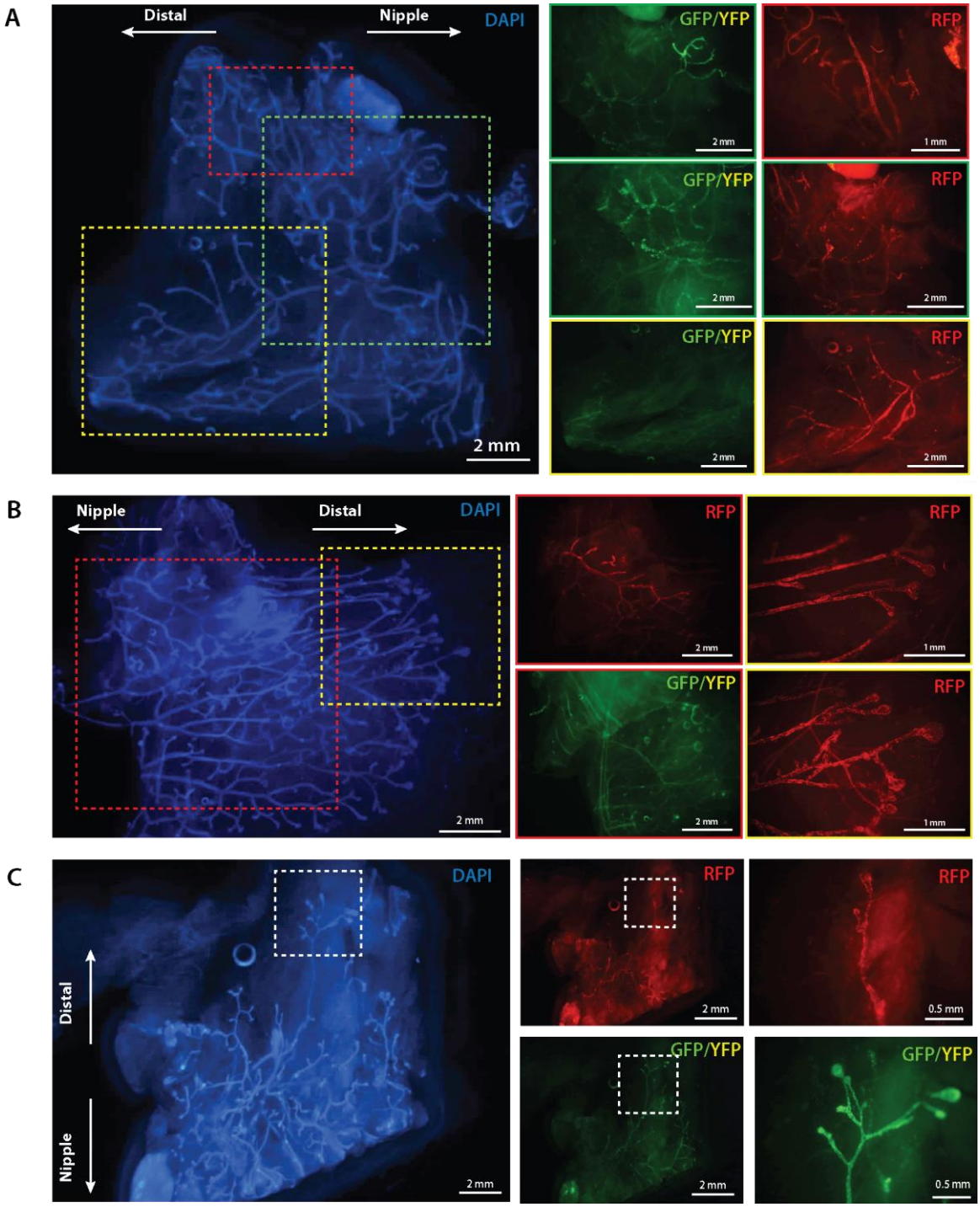


Figure 4.2.1.6 Wholemount fluorescence stereomicroscopy images of the mammary ductal network observed in 6 week old R26R-Confetti mice labelled in utero.

Pregnant dams were induced with tamoxifen via oral gavage at approximately E16.5-E17.5. Labelling was observed in the resulting hemizygous R26-Confetti offspring at 6 weeks of age. Representative examples of GFP. YFP and RFP labelling in mammary glands co-stained with DAPI nuclear dye are shown with zooms from different mice in **(A)**, **(B)**, and **(C)**, showing the extent of labelling, from the nipple region to outer edges of the fat pad.

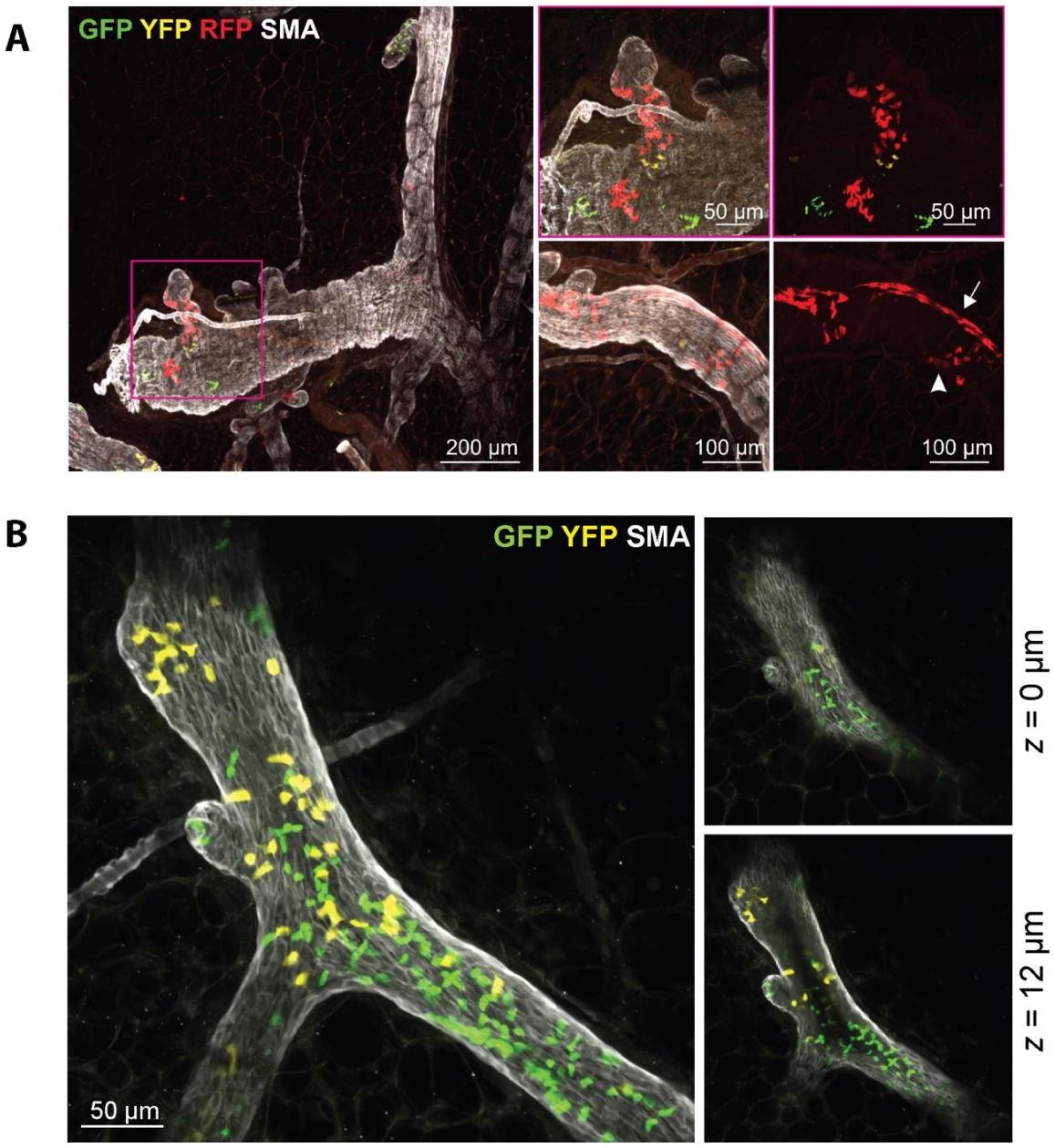


Figure 4.2.1.7 Examples of multi-colour regions in R26R-Confetti mice in mammary glands of 6 week old mice labelled in labelled in utero.

(A) Example of a multicoloured nipple region. Magnified regions (inset) show distinct clonal regions, in addition to presumed bipotent clones (arrow and arrowhead indicate adjacent RFP+ basal and luminal cells, respectively). **(B)** Representative example of a multicolour distal branch labelled with both luminal GFP and YFP clones and SMA co-staining. Z sections of the region with corresponding z-values are shown (inset).

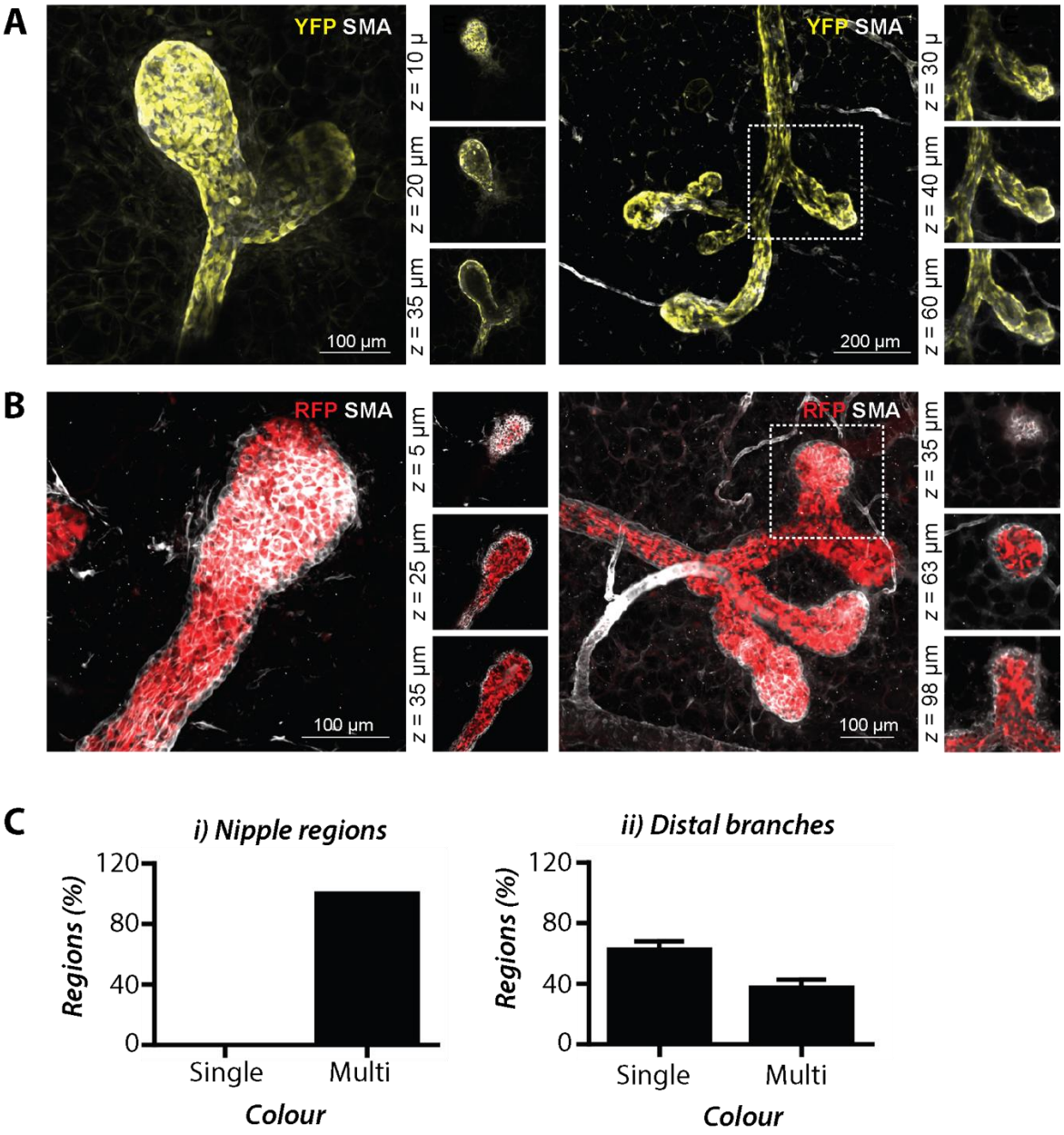


Figure 4.2.1.8 Examples of uni-colour labelled distal regions in mammary glands of 6 week old R26R-Confetti mice labelled in utero.

(A) Representative example of single-colour labelling in terminal end buds distal branches (left and right, respectively) comprising YFP+ basal cells. **(B)** Representative example of single-colour distal branches and TEBs comprising RFP+ luminal cells. Z sections of the regions with corresponding z-values are shown (inset). **(C)** Graphs showing the percentage of single- and multi-coloured nipple regions (i) and distal branches (ii). Data are mean±s.e.m. of 940 cells (from seven nipple regions, n=5 mice) and 4439 cells (from 85 distal branches, n=7 mice) from randomly-selected 3D images.

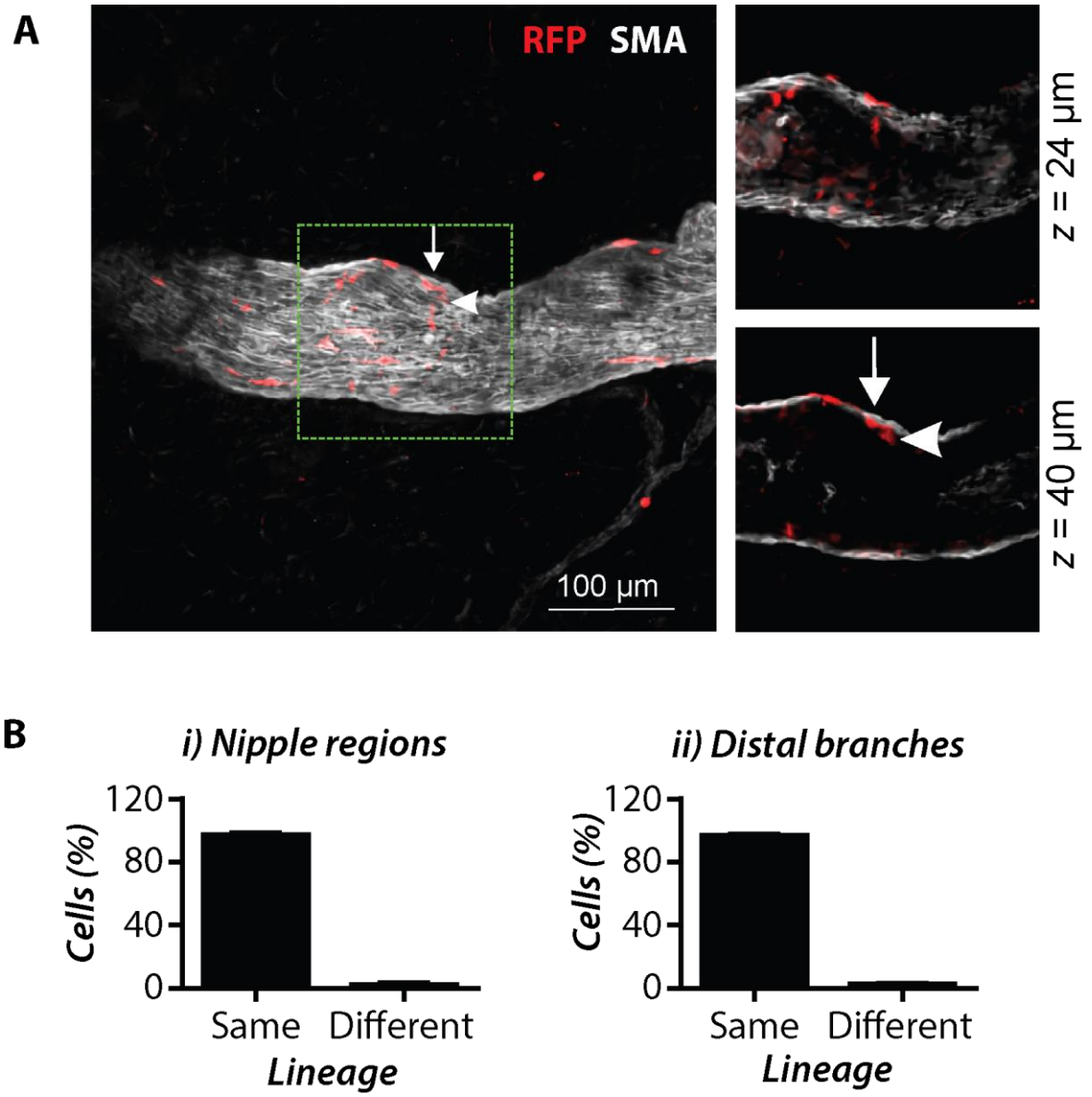


Figure 4.2.1.9 Labelled distal regions in mammary glands of 6 week old R26R-Confetti mice labelled in utero.

(A) Example of a rare distal branch containing interspersed RFP+ luminal and basal cells. Images show maximum-intensity z-projections and optical slices of a region of interest (boxed and enlarged in panels). Arrow and arrowhead show adjacent RFP+ basal and luminal cells. **(B)** Cell neighbour analysis nipple regions (i) and distal branches (ii), showing that the majority of labelled cells had a same-colour neighbour of the same lineage. Data are mean±s.e.m. of 940 cells (from seven nipple regions, n=5 mice) and 4439 cells (from 85 distal branches, n=7 mice) from randomly-selected 3D images.

4.2.2 Tracing pubertal mammary stem cells

We next wanted to examine the role of mammary stem and progenitor cells during pubertal development. Similarly to embryonic MaSCs, there are no definitive markers and as such, and to avoid bias, this again requires the use of neutral lineage tracing models. We therefore utilised the same models to study pubertal morphogenesis as to those used in embryonic tracing, and again combined with optical tissue clearing and 3D imaging.

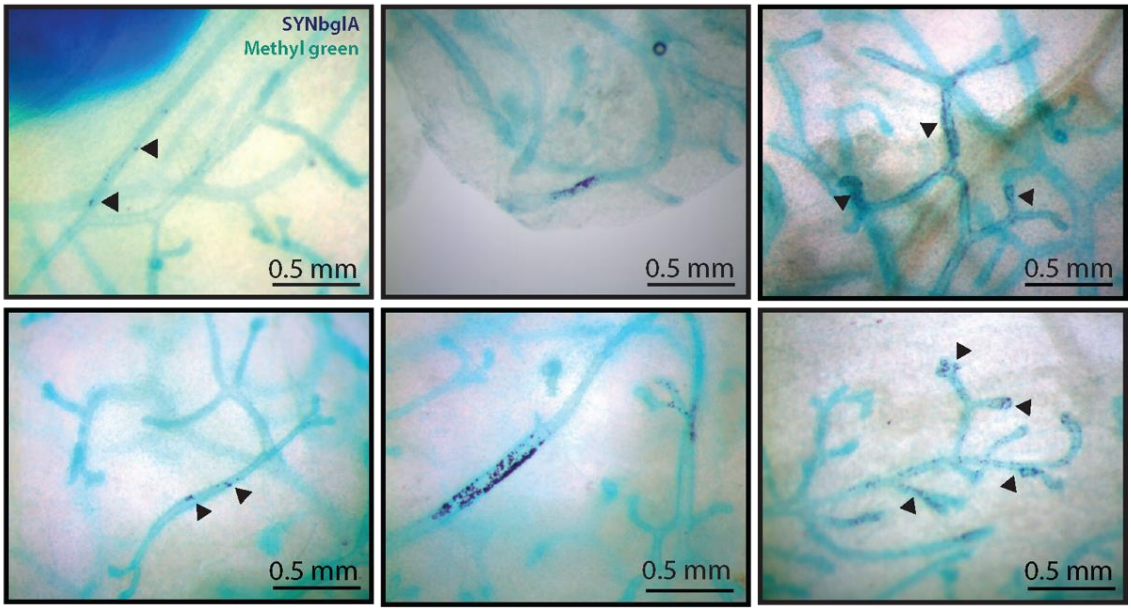
4.2.2.1 R26$^{[CA]30}$ SYNgblA

The first model we utilised to study pubertal mammary gland development was the R26$^{[CA]30}$ SYNbglA model. Use of the R26$^{[CA]30}$ SYNbglA model enabled macroscopic clonal analysis by wholemount histochemistry. Labelling was examined in tissues from the mammary glands of mice 7 weeks of age, towards the end of puberty. Ducts containing variable numbers of SYNbglA+ cells interspersed with unlabelled cells could be visualised *in situ*, with ductal structures visualised using methyl green counter staining and CUBIC optical clearing, as outlined in the previous chapter (Figure 4.2.2.1a).

In order to determine the slippage frequency in the pubertal mammary epithelium, clone abundance, size and distribution were examined when functionally active MaSCs are presumed to drive ductal elongation and branching morphogenesis. The amount of small (approximately 2-5 cells), moderate (approximately 5-50 cells) and large (> 50 cells) SYNbglA+ regions were quantified per gland (from n=9 mice, examples shown in Figure 4.2.2.1a). The majority of observed clonal regions were small to moderate, containing less than 50 cells (Figure 4.2.2.1b), presumably having arisen from proliferation-limited progenitors or terminally differentiated cells. On the other hand, large clones were much rarer, occurring at a rate of approximately 0.03 ± 0.06 per mammary gland. Based on the clone sizes, which sometimes was upwards of several hundred cells, and the number of branches that these regions occupied, it is likely that these arise from a single MaSC. Considering all regions, overall strand slippage was extremely rare with approximately 1.49 ± 0.92 events observed per gland – a similar figure to that seen in the intestine (Kozar et al., 2013) (Figure 4.2.2.1b). Consequently, the likelihood of clone convergence under these parameters is exceedingly low, and therefore provides confidence of true single cell lineage tracing using the R26$^{[CA]30}$ model.

In order to find more of these rarer slippage eve1nts we then analysed a further 240 mammary glands from 30 mice. To gather more information about the potential MaSC characteristics, we quantified the length and the number branches that the clones occupied, in addition to the length from the nipple as an inference of the time that slippage occurred (see above, summarised in Figure 4.2.2.2). In total, we observed a further 8 large clonal regions, spanning up to 9mm in length and contributing to 20 major branches (representative examples shown in Figure 4.2.2.3). Interestingly, a similar labelling pattern to that seen in the embryonic tracing was observed, with SYNbglA+ cells seen to be intermixed with unlabelled cells. This was likely due to the dilution of marked daughter cells with the progeny of unmarked MaSCs/progenitors during ductal expansion in puberty.

A Number of SYNbglA+ cells →

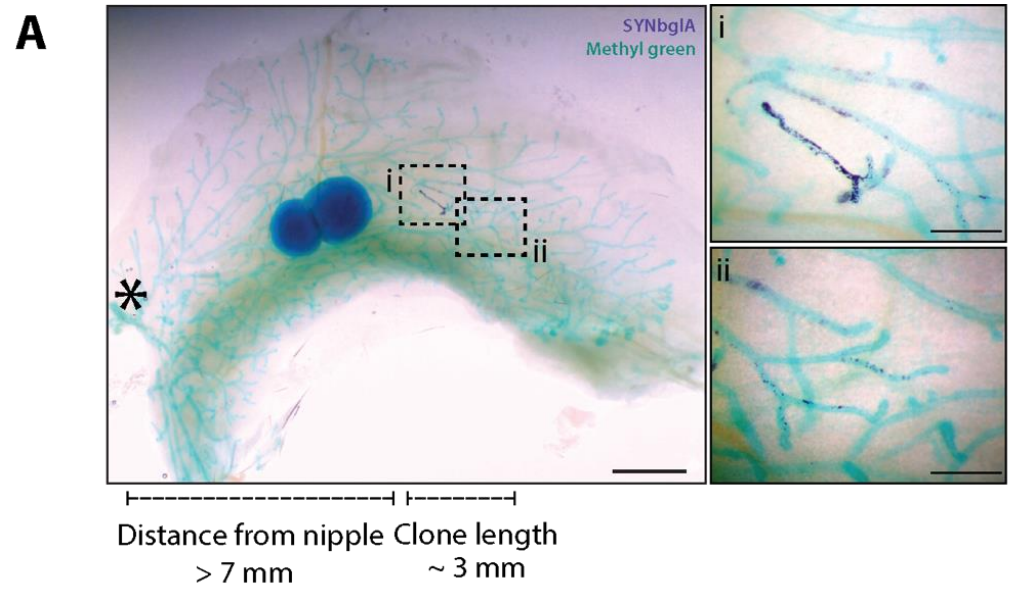


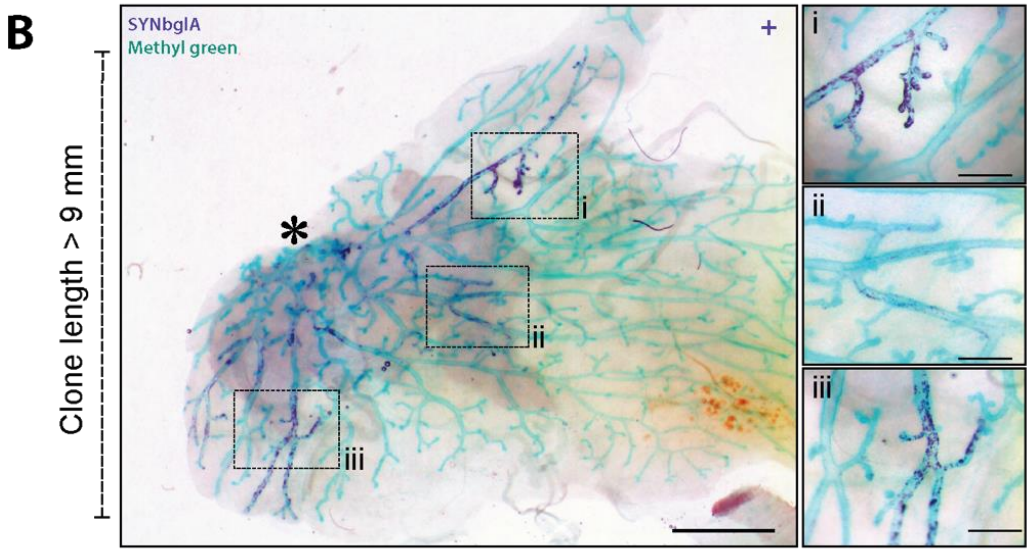
B

Clone size	Small	Moderate	Large
Average no. per gland	0.57 ± 0.69	0.82 ± 0.27	0.10 ± 0.08
Total		1.49 ± 0.92	

Figure 4.2.2.1 Calculating the rate of strand slippage in lineage tracing of R26$^{[CA]30}$ SYNbglA pubertal animals.

The event rate in the mammary gland was determined using R26$^{[CA]30}$ SYNbglA mice, as this model is conducive to macroscopic analysis of all labelling events. **(A)** Example of different sizes of SYNbglA clones in mammary glands of R26$^{[CA]30}$ SYNbglA mice observed at 7 weeks of age. Examples of small (approximately 2-5 cells), moderate (approximately 5-50 cells) and large (> 50 cells) labelled SYNbglA regions are shown, marked with arrow heads. **(B)** Quantification of labelling from of each of these bins, and the total rate of strand slippage, reflecting the chance of observing any labelling event. Note, due to the stochastic nature of labelling, it was associated with a large error. Values show mean ± S.D. quantified from 9 pubertal hemizygous mice.





C

	R26[CA]30 SYNbglA							
Total no.	30 mice (240 mammary glands)							
Linear length (mm)	> 8	~ 5	> 4	> 3	~ 3	~ 4	> 9	~ 3
No. branches	> 18	> 10	> 15	> 12	> 19	> 9	> 20	> 7
Distance from origin (mm)	> 3	> 1	> 5	> 1	> 1	> 3	0*	> 4

Figure 4.2.2.2 Examples of large clonal regions in R26[CA]30 SYNbglA pubertal animals.

(A) and **(B)** Representative examples of SYNbglA labelling, with length of clonal regions denoted and distance from nipple (marked with an asterisk). Magnified views are shown inset. **(C)** Summary of the eight large clonally marked regions (likely to have arisen from the labelling of a MaSC/progenitor) observed from the analysis of 30 R26[CA]30 SYNbglA mice.

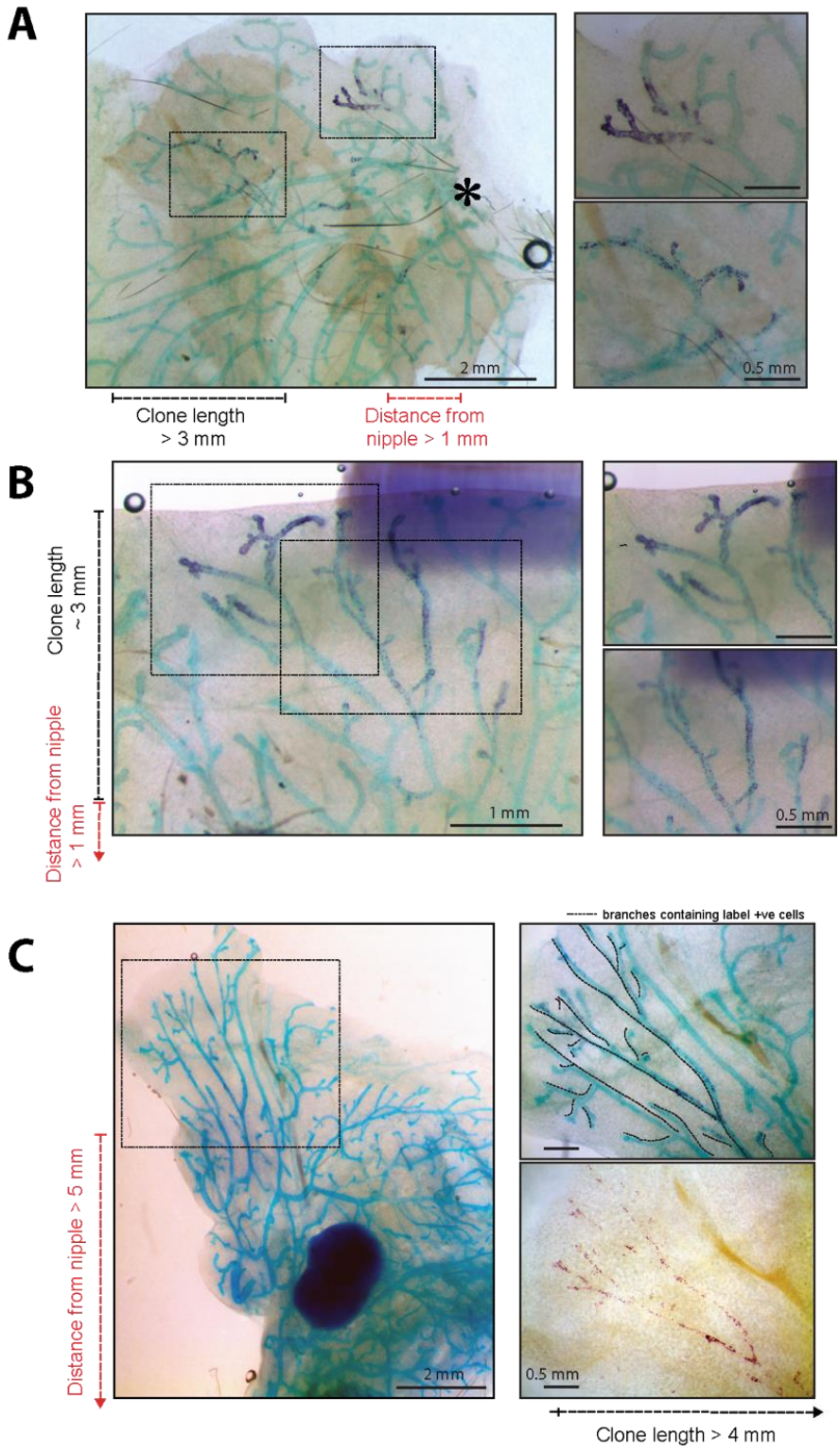


Figure 4.2.2.3 Further examples of large clonal regions in R26[CA]30 SYNbglA pubertal animals.

(A), **(B)** and **(C)** Representative examples of SYNbglA labelling in mammary galnds from different mice co-stained with Methyl Green epithelial counter stain. The length of clonal regions is denoted, with distance from nipple (marked with an asterisk). Magnified views are shown inset.

4.2.2.2 R26$^{[CA]30}$ YFP

The R26$^{[CA]30}$ SYNbglA model allowed for wholemount imaging and the identification of macroscopic clonal structures *in situ*, which is suitable for observing the entire gland in order to quantify labelling and observe patterns. However, for a more definitive determination of lineage on the basis of slippage co-localisation with markers of specific mammary gland cellular populations, a supplementary model is required. Consequently, the R26$^{[CA]30}$ YFP model was used. As with the R26$^{[CA]30}$ SYNbglA model, the reporter gene (in this case, YFP) becomes randomly expressed during cell replication (Figure 4.2.2.4a). If slippage occurs in a mammary stem or progenitor cell, this will then allow tracing of its progeny.

For examining mammary stem and progenitor cell involvement in pubertal development we again investigated labelling at 7 weeks of age in mice hemizygous for R26$^{[CA]30}$ YFP. SeeDB-based optical tissue clearing was used for 3D visualisation of the ductal network, combined with immunolabelling for markers of basal (SMA) and luminal (K8) cell lineages. Accordingly, the progeny arising from a single fluorescently-marked cell could then be visualised using confocal imaging and characterised *in situ* with single cell resolution. Clonal expansion of YFP+ cells produced a distinctive mosaic labelling pattern, with both luminal and basal clonal regions detected (Figure 4.2.2.4b, left and right respectively). Importantly, these matched patterns observed in the R26$^{[CA]30}$SYNbglA model. Together, these data strongly support the hypothesis that more than one MaSC/progenitor also contributes to the elongation of each major duct during puberty.

In order to further characterise mammary stem or progenitor cell involvement in pubertal ductal expansion we analysed a further 504 mammary glands from a total of 63 mice. Similar to what was seen in the R26$^{[CA]30}$ SYNbglA model, clones differed in size with the majority containing only a few cells. Despite the vast number of mammary glands examined, only three large clonal regions were observed (assumed to be arisen from a true MaSC due to the large number of cells involved) (summarised in Figure 4.2.2.4c). These appeared to all be of single lineage, but we next sought to further confirm this and characterise these clones in terms of composition and contribution to ducts. Since the timing of the slippage event cannot be determined in the R26$^{[CA]30}$ model, it was again assumed that if clones were observed 1 mm proximally from the nipple region then slippage had occurred postnatally (Paine et al., 2016). Glands were imaged using stereomicroscopy to allow clones to be visualised in the context of

the entire ductal network. Higher resolution 3D confocal images of the clonal regions, which allow for single cell resolution, were then mapped back onto these stereomicroscopy images.

The first large clonal region examined spanned over 7 mm in length and was comprised of over 10 major branches. Importantly, it occurred more than 2 mm past the nipple region (Figure 4.2.2.5a, marked with an asterisk), suggesting labelled cells were the progeny of a postnatally slipped cell. Upon 3D visualisation of the YFP+ cells, the lineage of all labelled cells, as determined by cell location in the duct and co-localisation with lineage markers, was luminal (Figure 4.2.2.5b), demonstrating that the cells are the progeny of a lineage restricted stem cell.

We next observed a clone which spanned an even more vast region of the gland, spanning over 8 mm in length and occupying more than 20 branches when visualised using stereomicroscopy (Figure 4.2.2.6a and Figure 4.2.2.7a). Upon closer inspection of this clone with confocal microscopy (Figure 4.2.2.6b), labelled cells were seen on both major branches (for example Figure 4.2.2.6a,iii and Figure 4.2.2.6c,i,iv) and in terminal end ducts and side branches (Figure 4.2.2.6b,i and c,vi). By morphology and co-localisation with K8, slipped cells all appeared to be of the luminal lineage. However, due to the vast nature of this clone supplementary confirmation was required.

Consequently, to further investigate the adult MaSCs/progenitors and quantify their contribution to ductal morphogenesis, we adopted the use of imaging algorithms in collaboration with Dr Leila Muresan (see methods for detailed procedure). Briefly, Gaussian filtering was used on the entire 3D stacks to exclude noise and as such identify YFP+ cells within the ductal structures only. Labelled cells were then segmented and classified as either luminal or basal depending on the intensity values of the voxels in either the K8 or SMA channel; if the number of K8 voxels exceeded the number of SMA voxels then that cell would be classified as luminal and *vice versa*.

By 3D volumetric segmentation, all cells analysed using this method expressed K8. For further confirmation, the gland was sectioned over its entire depth and traditional histochemical analysis carried out using fluorescent co-staining with YFP (Figure 4.2.2.9b). Lineage restriction was confirmed over the 300µm depth, with all YFP+ cells co-localising with K8 (Figure 4.2.2.9c). These data support the observations with the $R26^{[CA]30}$ SYNbglA mice, providing

further evidence that unipotent MaSCs/progenitors contribute extensively to ductal morphogenesis.

Of note, whilst analysing the vast number of glands we noticed the differential expression of K8 in luminal cells, identifying what appeared to be two distinct populations comprising of $K8^{lo}$ and $K8^{hi}$ cells (Figure 4.2.2.9a). In order to characterise these further we therefore investigated the co-expression of hormone receptors of these cells (Figure 4.2.2.10a). Co-staining with progesterone receptor (PR) displayed $K8^{hi}$ cells co-express PR, whereas $K8^{lo}$ cells generally do not (Figure 4.2.2.10b). Interestingly, both populations of these cells appear to be present in a similar proportion to one another in the virgin mammary gland.

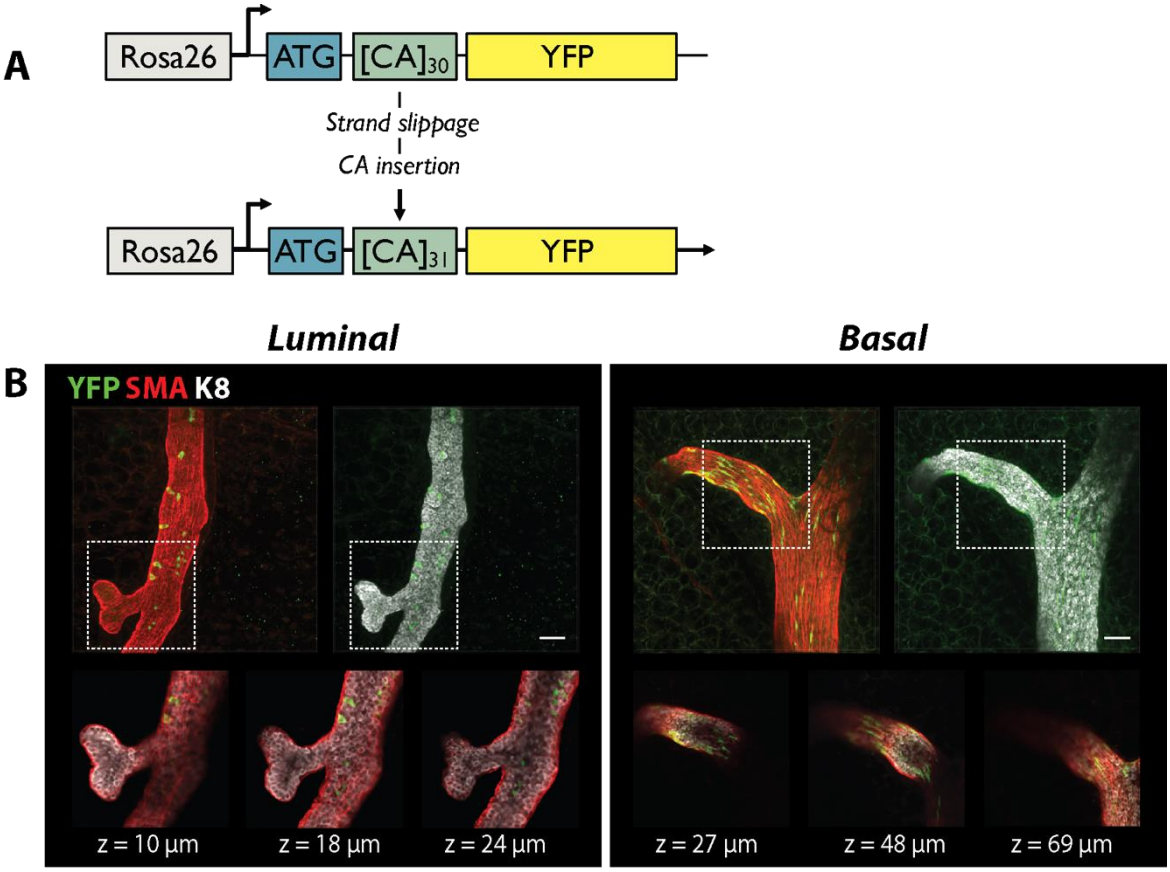
In a number of cases, labelled cells appeared to encompass both these populations of luminal cells (for example in Figure 4.2.2.9a, arrow heads). Therefore, a custom computational algorithm was developed to determine whether a single marked luminal mammary stem/progenitor cell contributed equally to both $K8^{hi}/PR^{+}$ and $K8^{lo}/PR^{-}$ populations. In each of the four cases analysed within a single clonal region, this revealed that the K8 signal inside the YFP+ cells were significantly higher than the total luminal population, with a modest but significant overrepresentation of the YFP label in $K8^{hi}$ cells; thus the null hypothesis of equality of the two distributions was rejected (Figure 4.2.2.10c, *Kolmogorov-Smirnov* test at $p < 0.05$).

We next developed another computational algorithm to determine the potential contribution of an active MaSC/progenitor to each duct. To do so, a volumetric ratio of YFP+ cells relative to the total cellular volume of the ductal structure was computed (Figure 4.2.2.11a). This calculated that YFP+ cells comprised on average 4.7 ± 1.7% of the total cellular volume in this region, demonstrating the contribution of a single parent MaSC (Figure 4.2.2.11b). Assuming that other unlabelled MaSCs have similar proliferative and competitive capacities, these data suggests that there may be at least 20 lineage-restricted luminal MaSCs within each major duct that produce the luminal cells and drive ductal morphogenesis during puberty.

Interestingly, only one large YFP+ basal clone was observed out of the 63 animals examined. This clone was slightly smaller than the luminal clones identified, comprising of approximately 4mm of the pubertal gland and occupying 9 branches (Figure 4.2.2.12a). Following examination of the clone using 3D confocal imaging, labelled cells followed the same non-contiguous

pattern as the luminal clones (Figure 4.2.2.12b). To confirm the lineage restriction of this clone, we carried out similar computational analysis, with minor alterations.

The different morphology of basal cells compared with luminal cells, together with the increased noise due to the depth of the clone, necessitated adaption of the algorithm: firstly, a denoising and deblurring step was applied to increase the signal-to-noise ratio of the images. Following this, a seeded watershed was used whereby the ductal structures were inverted and segmented to allow clear visualisation of the bi-layer, in addition to better co-localisation. As expected, this computed that all 53 cells examined co-localised with SMA and as such were of the basal lineage (Figure 4.2.2.12d). A similar volumetric analysis was also carried out suggesting that, on average, clonal regions occupied 5.8 ± 3.2% of the ducal regions analysed (Figure 4.2.2.12c). Assuming other basal stem and progenitors cells possess a similar capacity, this suggests that there could be at least 15 MaSCs that contributed to the basal lineage in this region – slightly lower than the luminal prediction, but this could be attributed to the reduced total number of basal cells in comparison to luminal cells in the mouse mammary gland (Rios et al., 2016).



C

	R26[CA]30 YFP		
Total no.	63 mice (504 mammary glands)		
Linear length (mm)	> 8	> 7	> 4
No. branches	> 20	> 10	> 9
Distance from origin (mm)	> 2	> 2	> 2
Composition (Lu/Ba/Mixed)	Lu	Lu	Ba

Figure 4.2.2.4 Overview of the R26[CA]30 YFP mouse model for pubertal lineage tracing

(A) Schematic representation of the R26[CA]30 YFP construct. **(B)** Representative examples of luminal and basal (left and right panels, respectively) YFP+ clonal regions from mice, representing over 25 label-positive regions. Z sections of the regions with corresponding z-values are shown (inset), scale bars, 50 mm. **(C)** Summary table of the three clonally marked regions seen in pubertal R26[CA]30 YFP mammary glands, presumed to have arisen from the labelling of a MaSC/progenitor, from the analysis of 504 mammary glands from 63 R26[CA]30 YFP mice.

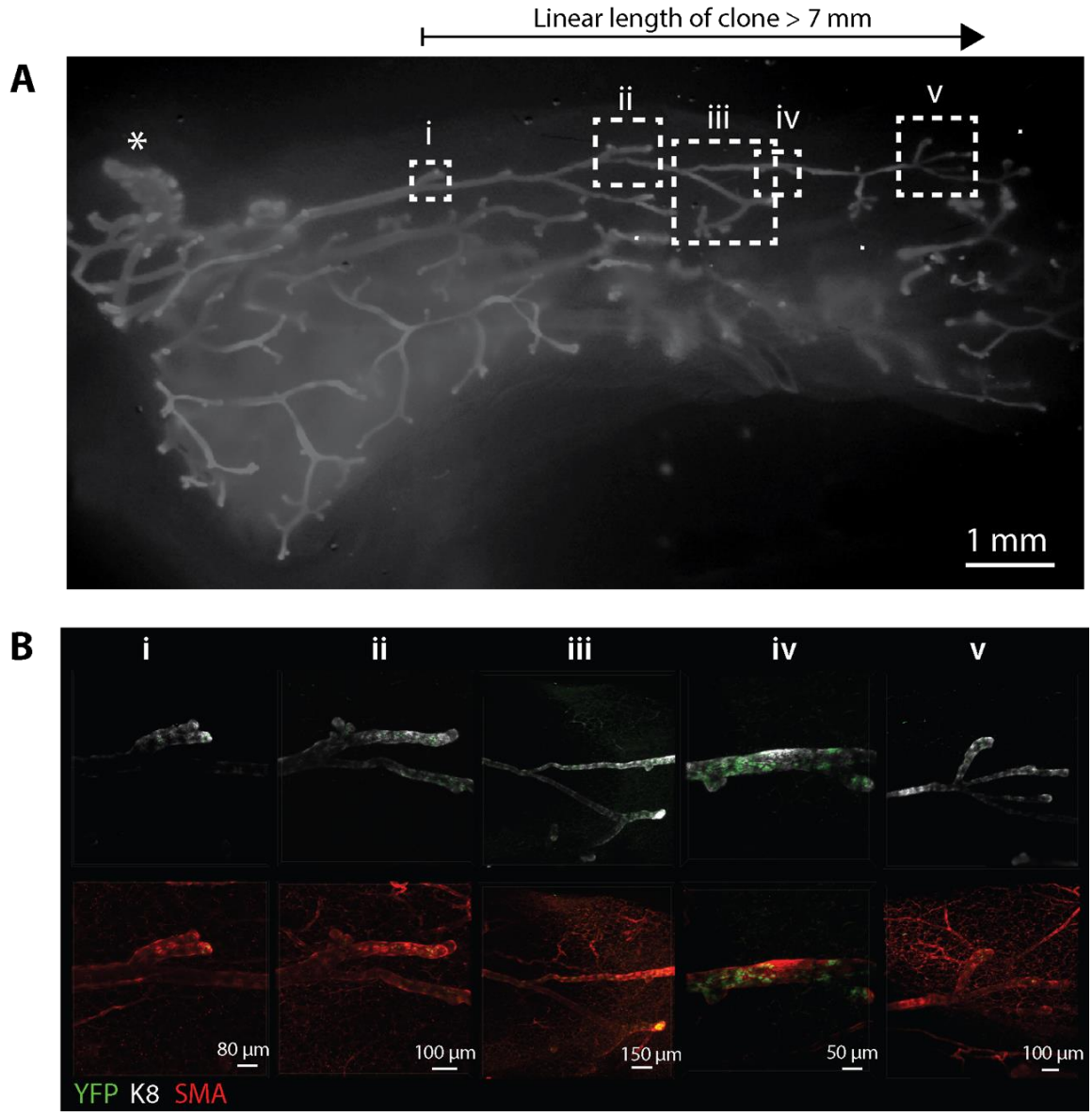


Figure 4.2.2.5 Example of clonal labelling in R26$^{[CA]30}$ YFP mouse model during pubertal lineage tracing

(A) Wholemount stereomicroscopic images of K8 immunofluorescence, showing the mammary ductal network of a 7 week R26$^{[CA]30}$ YFP mouse, demarcating the linear length of the clone and regions that were imaged at high cellular resolution by confocal microscopy shown in **(B)**, magnified views i–v. Asterisk shows the location of the nipple.

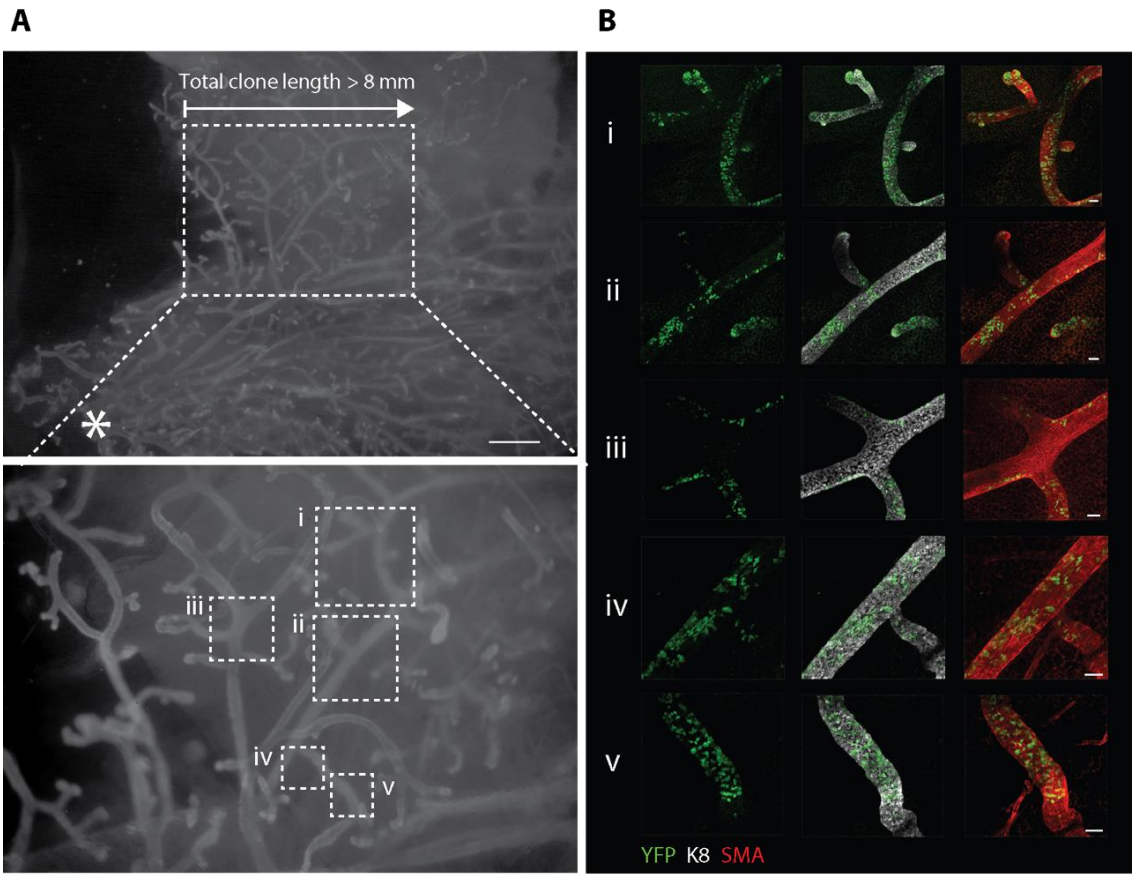


Figure 4.2.2.6 Further example of clonal labelling in R26$^{[CA]30}$ YFP mouse model during pubertal lineage tracing

(A) Wholemount stereomicroscopic images of K8 immunofluorescence, showing the mammary ductal network of a 7 week R26$^{[CA]30}$ YFP mouse, demarcating the linear length of the clone and regions that were imaged at high cellular resolution by confocal microscopy shown in **(B)**, magnified views i–v. Asterisk shows the location of the nipple. Scale bar, 1mm (wholemount) and 50µm (confocal).

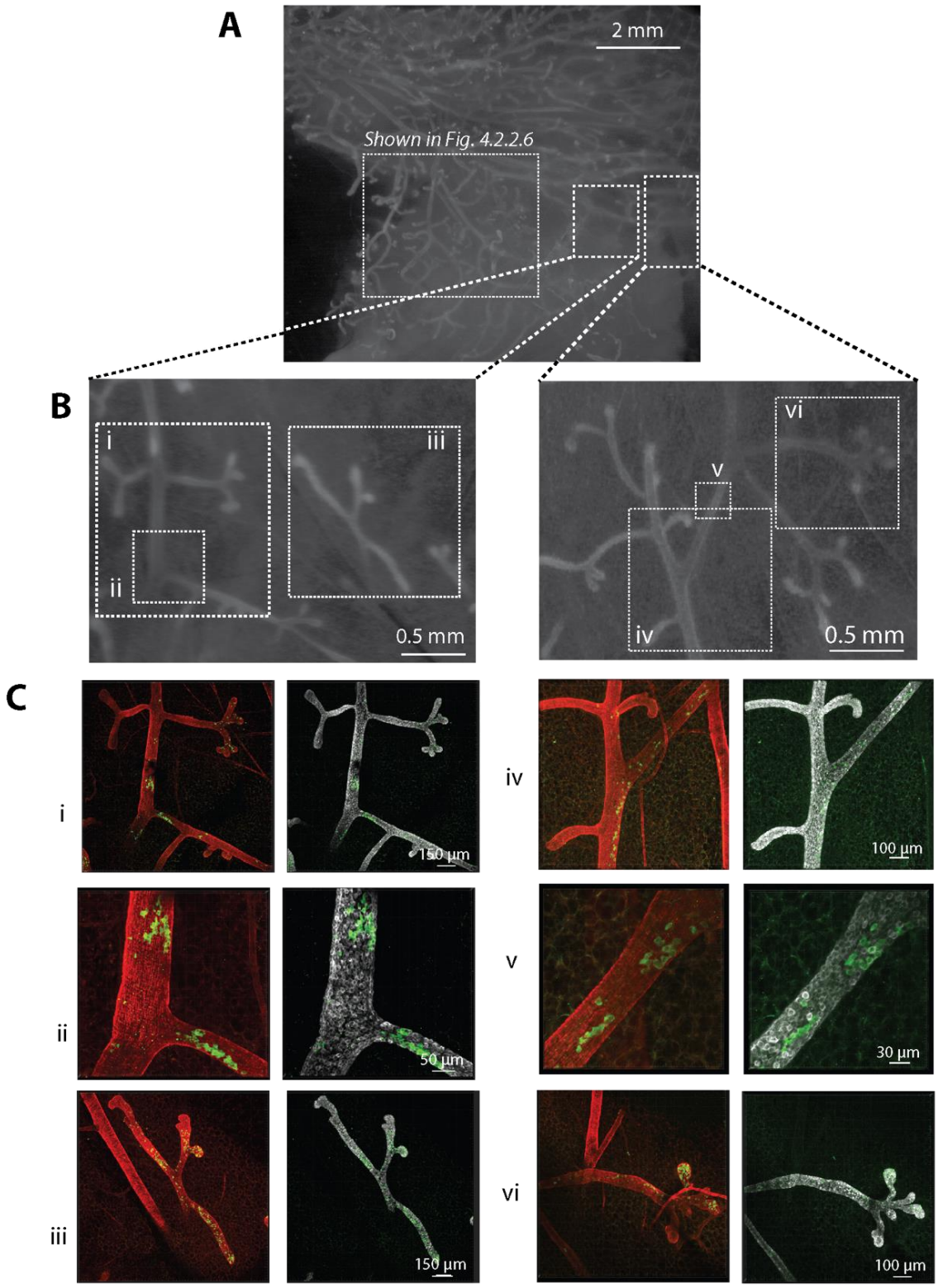


Figure 4.2.2.7 Further views of labelling in R26$^{[CA]30}$ YFP mouse mammary gland shown in Figure 4.2.1.6

(A) Wholemount stereomicroscopic images of K8 immunofluorescence, showing the mammary ductal network of a 7 week R26$^{[CA]30}$ YFP mouse, demarcating the region shown in Figure 4.2.1.6 and magnified regions shown in **(B)** with further regions of interest that were imaged at high cellular resolution by confocal microscopy shown in (c), magnified views i–vi.

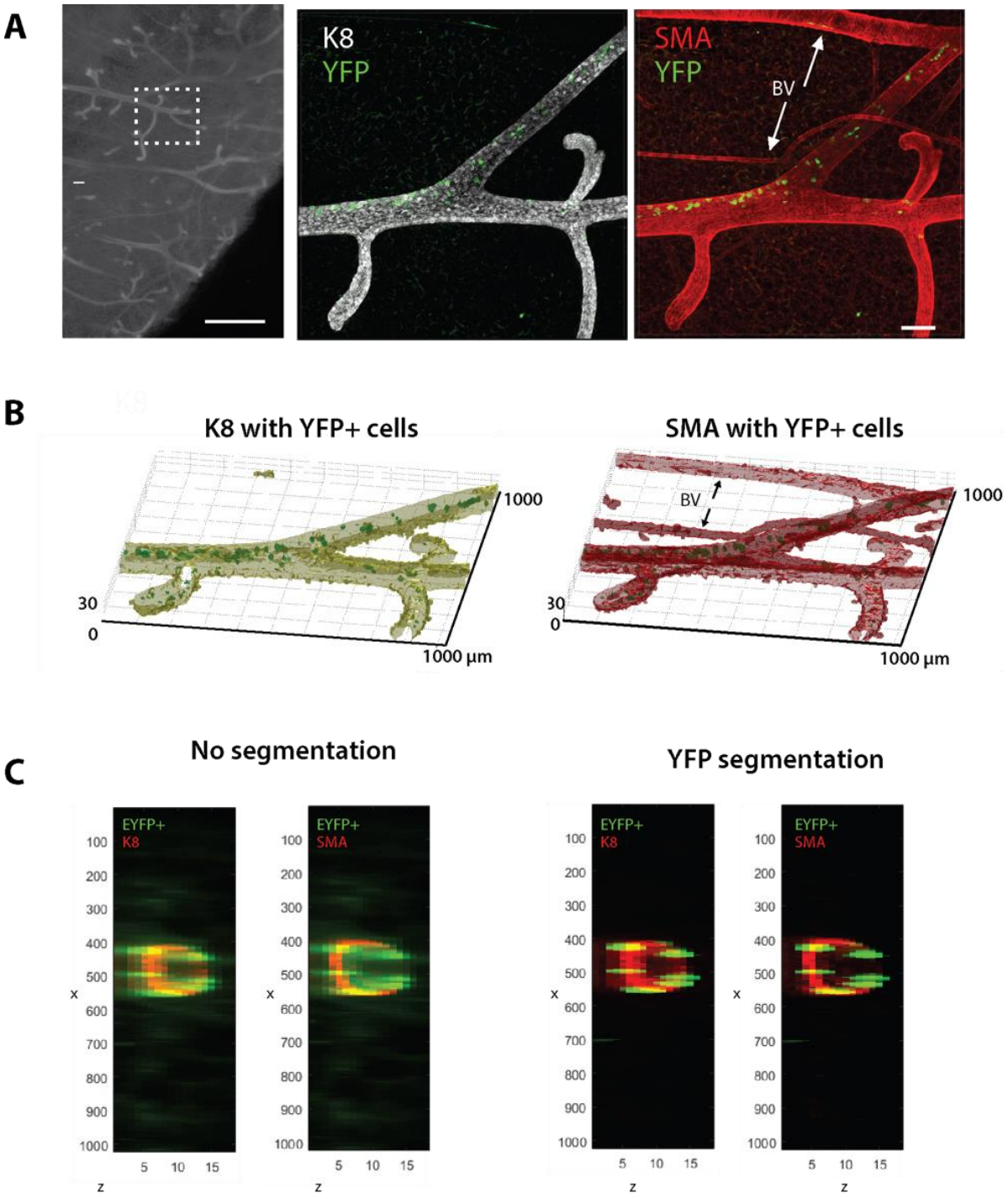
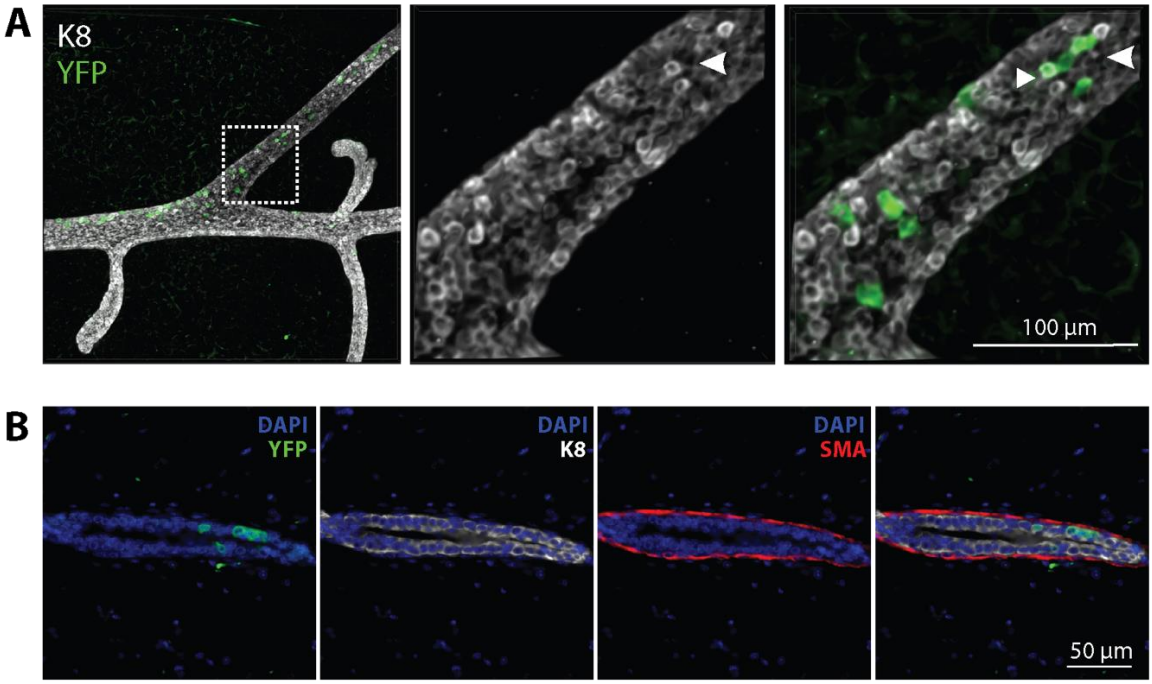


Figure 4.2.2.8. Computational analysis of clonal labelling in pubertal R26$^{[CA]30}$ YFP mammary gland shown in Figure 4.2.1.6& 7

(A) Images of a clonally marked region that was analysed by 3D image analysis, from fluorescent stereomicroscopy (left) and high-resolution 3D confocal microscopy. Scale bars, 1 mm (wholemount) and 100 µm (confocal). **(B)** and **(C)** Digital segmentation of YFP cells within the luminal (K8-expressing) and basal (SMA-expressing) compartments (left and right, respectively). BV, blood vessel. Algorithms developed in collaboration with Dr Leila Muresen.

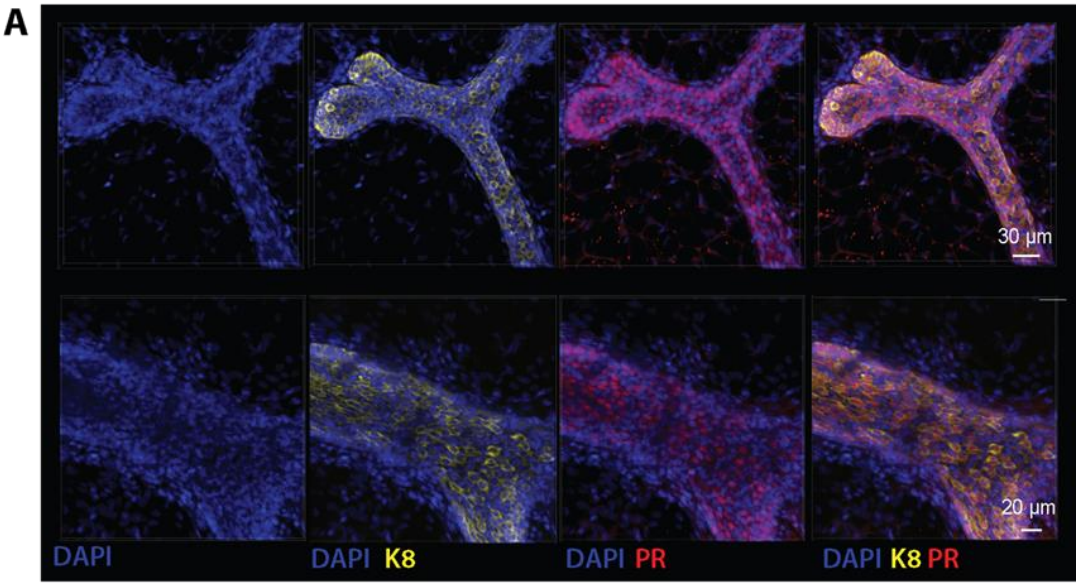


C

	YFP+ cells (2D analysis)	YFP events (3D analysis)
Luminal	281/281 (100%)	227/227 (100%)
Basal	0/281 (0%)	0/227 (0%)

Figure 4.2.2.9 Analysis of differential K8 expression in luminal cells in R26$^{[CA]30}$ YFP mice

(A) Example of K8^{hi} (arrow heads) and K8^{lo} cells seen in pubertal mammary glands, showing that progeny from a single luminal MaSC/progenitor encompass both K8^{hi} (arrowhead) and K8^{lo} luminal cell populations. **(B)** Fluorescent 2D immunohistochemical staining of a clonal region for lineage confirmation **(C)** Tabulated results of this lineage analysis using 2D and 3D analyses (left and right, respectively), confirming all YFP+ were luminal. For 3D analysis, all segmented ductal YFP cells were classified as luminal based on the proportion of K8 versus SMA signal (n=227 cells from 4 image sequences). For 2D analysis, cells were classified by manual scoring of histological sections (n=281 cells from 10 sections spanning 300m depth).



B

	Average ± S.D.
PR^{+} cells that are $K8^{hi}$	79.95 ± 8.10%

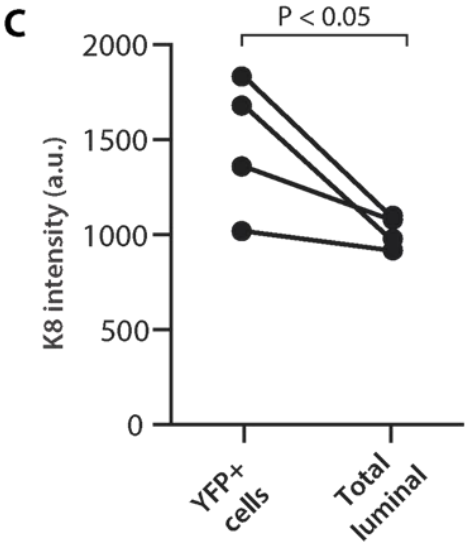
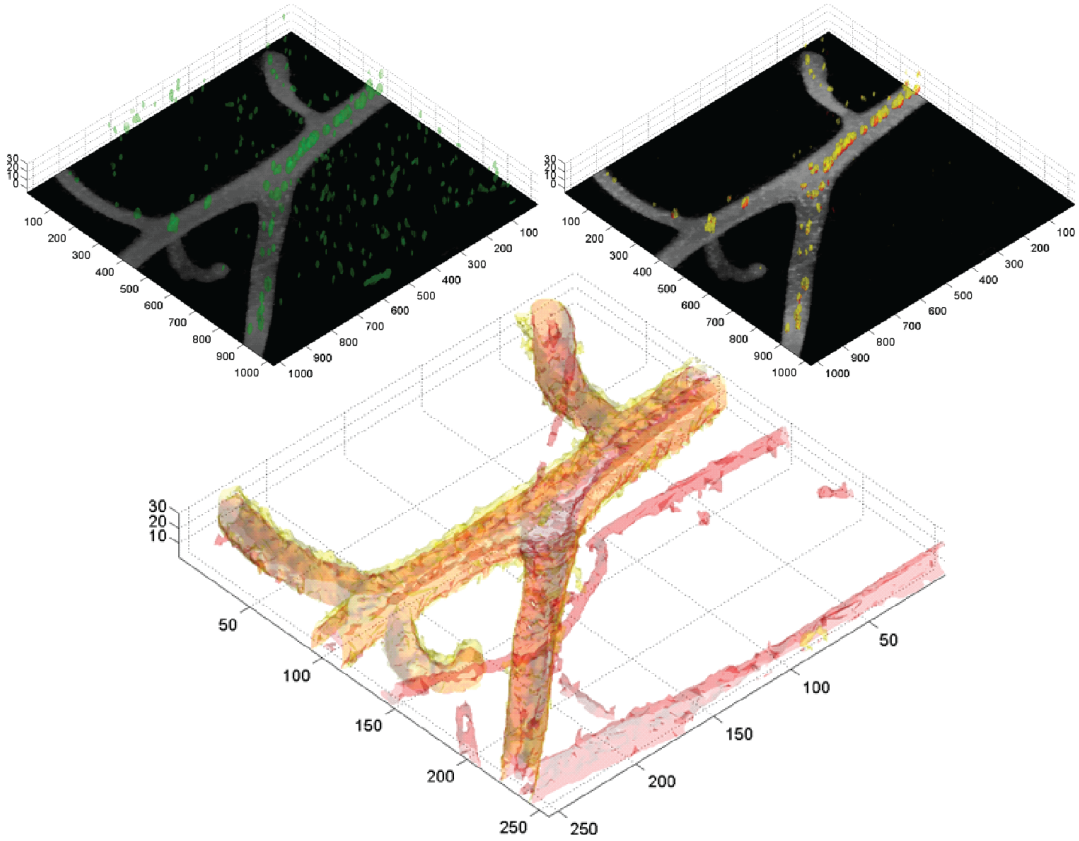


Figure 4.2. 2.10 Further analysis of differential K8 expression in luminal cells in R26$^{[CA]30}$ YFP mice

(A) Representative 3D confocal images showing non-uniform expression of K8 in luminal cells, with K8hi cells co-staining with nuclear progesterone receptor (PR). **(B)** Tabulated results, showing the proportion of PR positive cells scored as $K8^{hi}$ in virgin pubertal mammary glands. **(C)** Results from analysis to determine whether YFP labelled cells encompassed both K8 populations. A Kolmogorov-Smirnov test revealed that the K8 signal inside the YFP+ cells was significantly higher than the total luminal population. Data representative of the quantification of 6 images from 3 mice, with 1831 cells scored in total.

A

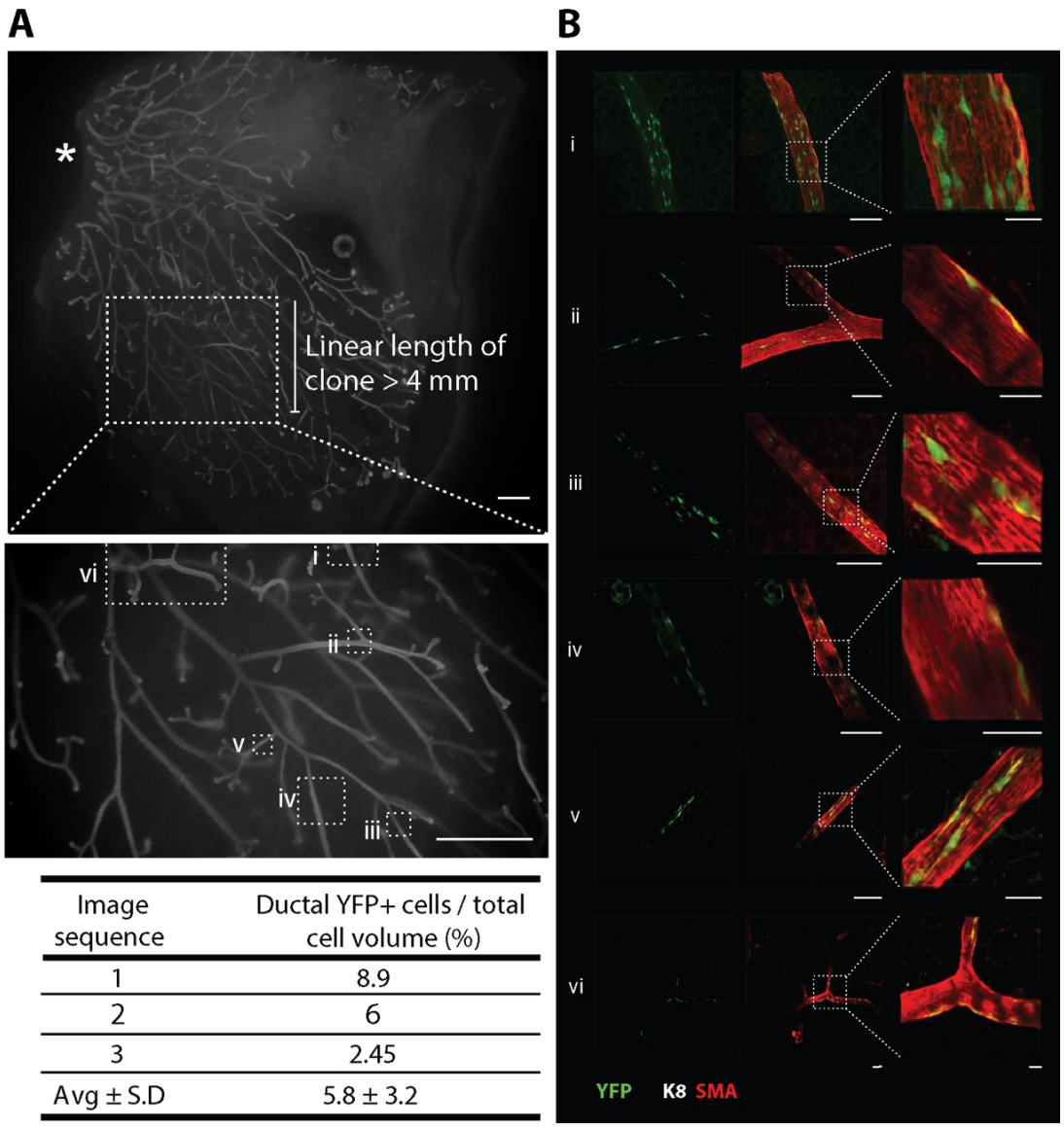


B

Image sequence	Ductal YFP+ cells / total cellular volume (%)
1	4.08
2	2.55
3	6.38
4	5.66
Avg ± S.D.	4.7 ± 1. 7

Figure 4.2.2.11. Volumetric computational analysis of clonal labelling in pubertal R26$^{[CA]30}$ YFP mammary gland shown in Figure 4.2.1.6&7

(A) Images were segmented and the volumetric contribution of the clone to the total ductal volume was computed, tubulated results shown in **(B)**. Analysis on four randomly selected regions.



C

Image sequence	Ductal YFP+ cells / total cell volume (%)
1	8.9
2	6
3	2.45
Avg ± S.D	5.8 ± 3.2

D

	YFP+ (3D analysis)
Luminal	0/53 (0%)
Basal	53/53 (100%)

Figure 4.2.2.12. Basal clonal labelling in pubertal R26$^{[CA]30}$ YFP mammary gland

(A) Wholemount stereomicroscopic images of K8 immunofluorescence, showing the mammary ductal network of a 7 week R26$^{[CA]30}$ YFP mouse, demarcating the linear length of the clone and asterisk showing the location of the nipple. Magnified area shows where the clone was detected (bottom panel) with areas (i-vi) that were imaged at high cellular resolution by confocal microscopy shown in **(B)**. **(C)** Tabulated results of the computed volumetric ratio of YFP cells, with respect to total basal cellular volume for each of the three regions analysed. **(D)** Tabulated results of the 3D lineage analysis, confirming the basal lineage of the clone (n=453 cells from 3 image sequences).

4.2.2.3 R26-Confetti

Although the R26$^{[CA]30}$ models provided important insights into the contribution of a single functional stem or progenitor cell to mammary gland development, we sought to confirm our observations with a second neutral approach that also allows the timing of the labelling event to be controlled and is also not dependent on a cell being in cycle at the time of labelling. We therefore used the R26-Confetti system again, this time in pubertal morphogenesis.

To achieve this, mice that were hemizygous for both R26-Confetti (Livet et al., 2007) and R26-CreER^{T2} (Ventura et al., 2007) were generated. To achieve the expression of the multicolour Confetti reporter, administration of tamoxifen is required for Cre-mediated recombination, which is achieved in a dose responsive manner (Livet et al., 2007; Ventura et al., 2007). As such, selection of the correct dose of tamoxifen is vital to accurately and confidently label single cells, whilst balancing this with a sufficient level of labelling. A titration of doses was carried out in order to select for optimal dose for single cell tracing. For this, a single dose of either 0.25mg, 0.5mg or 1mg of tamoxifen per mouse was injected via I.P injection at 4 weeks of age (Figure 4.2.2.13a). SeeDB- and CUBIC R1a-based clearing and 3D confocal imaging was utilised to visualise the clones and the amount of labelling was measured. To identify luminal and basal cells, wholemount immunostaining was performed using either K8 or SMA, respectively, and tissues were counterstained with DAPI to localise the mammary epithelium. Importantly, no labelling was observed in the vehicle injected control (Figure 4.2.2.13b)

All doses examined resulted in relatively low-density multicolour labelling in the mammary epithelium, allowing individual clones to be distinguished. Using this approach, we were able to visualise luminal and basal lineage-restricted GFP+, YFP+ and RFP+ clones, with the amount of labelling increasing in a dose responsive manner. CFP-expressing clones were rare and under-represented and were therefore not analysed (Figure 4.2.2.14); the reason for this is not entirely clear but may relate to the poor penetration of short wavelength light through thick specimens and the fine membranous localisation of the CFP reporter protein, which may be particularly sensitive to the optical clearing process. Alternatively, it may simply be a characteristic of the model, with differential expression of the confetti colours being previously identified in the mammary gland, prostate and pancreas (Wuidart et al., 2016; Sznurkowska et al., 2018).

To examine initial labelling, we examined glands 2 days following induction of lineage tracing with tamoxifen. Small regions of labelled cells were identified, mostly comprised of single cells (Figure 4.2.2.13c). Next, we examined glands following a 3 week chase to trace the progeny of cells labelled at the onset of puberty. In doing so, multiple clonal regions were observed. Importantly, labelling in the confetti mice was similar to that in the $R26^{[CA]30}$ models, in terms of both patterning and single clonal lineages, thus validating both approaches. However, a number of regions contained multiple clones of different colours (Figure 4.2.2.15a). In order to ensure confidence in tracing of single cells and labelling at a lower density to prevent clone convergence we proceeded to focus further analyses using 0.5mg per mouse (~35 μg/g body weight) of tamoxifen for low density lineage tracing.

At this dose, mammary branches labelled with single colours were observed (Figure 4.2.2.16a), in addition to branches with more than one colour (Figure 4.2.2.16b). To quantify labelling, the number of single- and multi-colour branches was calculated, which showed a chance of at least 50% clone convergence even at this low dose (Figure 4.2.2.16c). To characterise the lineages of the clones, cell neighbour analysis was carried out (as in the embryonic tracing, summarised in Figure 4.2.2.17a). This revealed that the majority of labelled cells had a same-colour labelled neighbour that was of the same lineage (Figure 4.2.2.17b, n=1419 cells distributed across 130 branches from 5 mice). Despite the higher chance of clone convergence, this supports the previous conclusions from the $R26^{[CA]30}$ models – namely that the mammary stem and progenitor cells involved in ductal morphogenesis are unipotent.

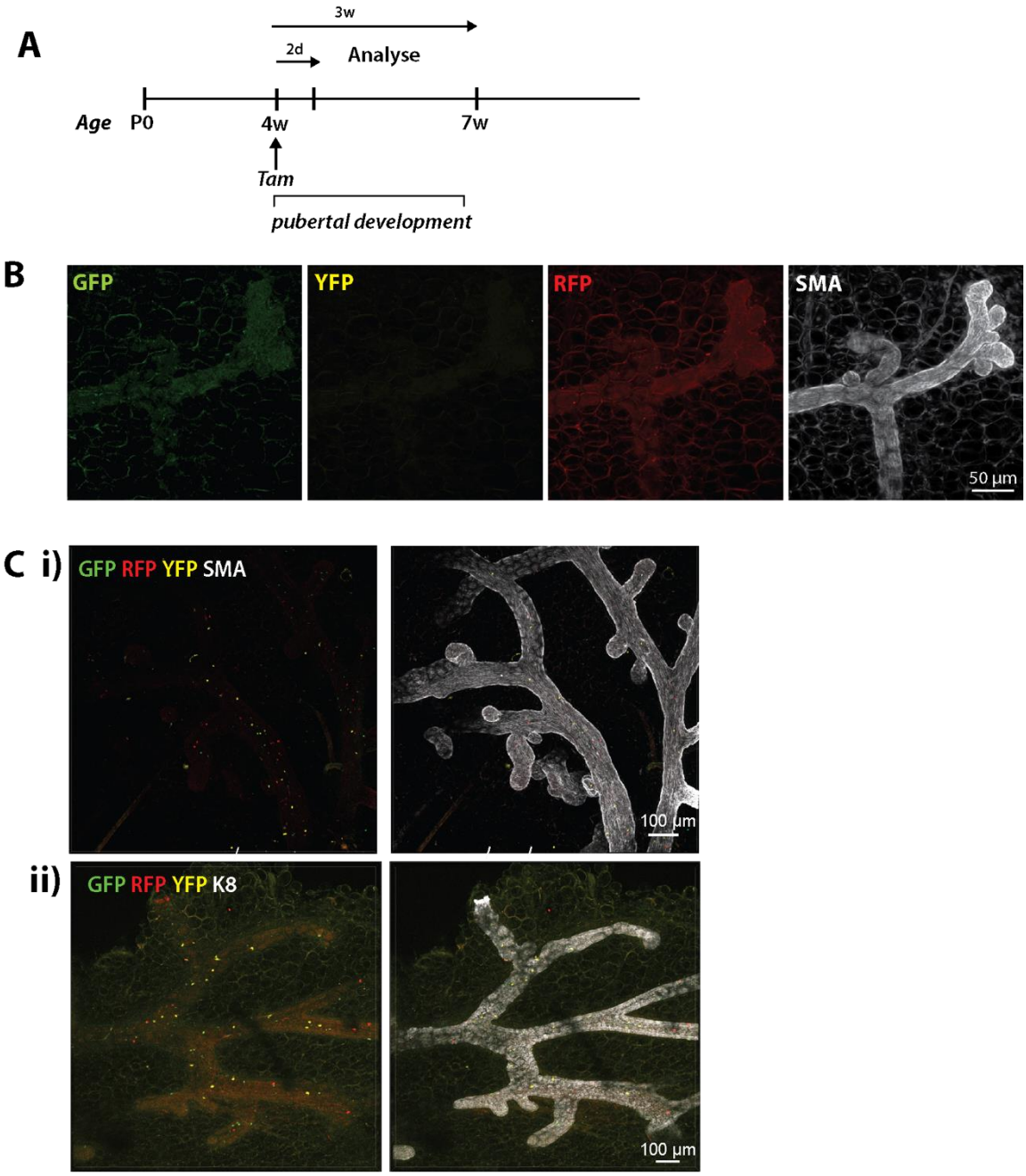


Figure 4.2.2.13. Overview of the use of the R26R-Confetti model for lineage tracing during branching morphogenesis

(A) Schematic of lineage tracing using the R26R-Confetti model during branching morphogenesis. Tamoxifen was administered at the beginning of puberty (4 weeks) and tissue harvested 3 weeks later (at 7 weeks). **(B)** showing the absence of labelling in control mice injected with oil (vehicle) **(Ci,ii)** Representative example maximum-intensity z-projections showing labelling following a short (2 day) chase.

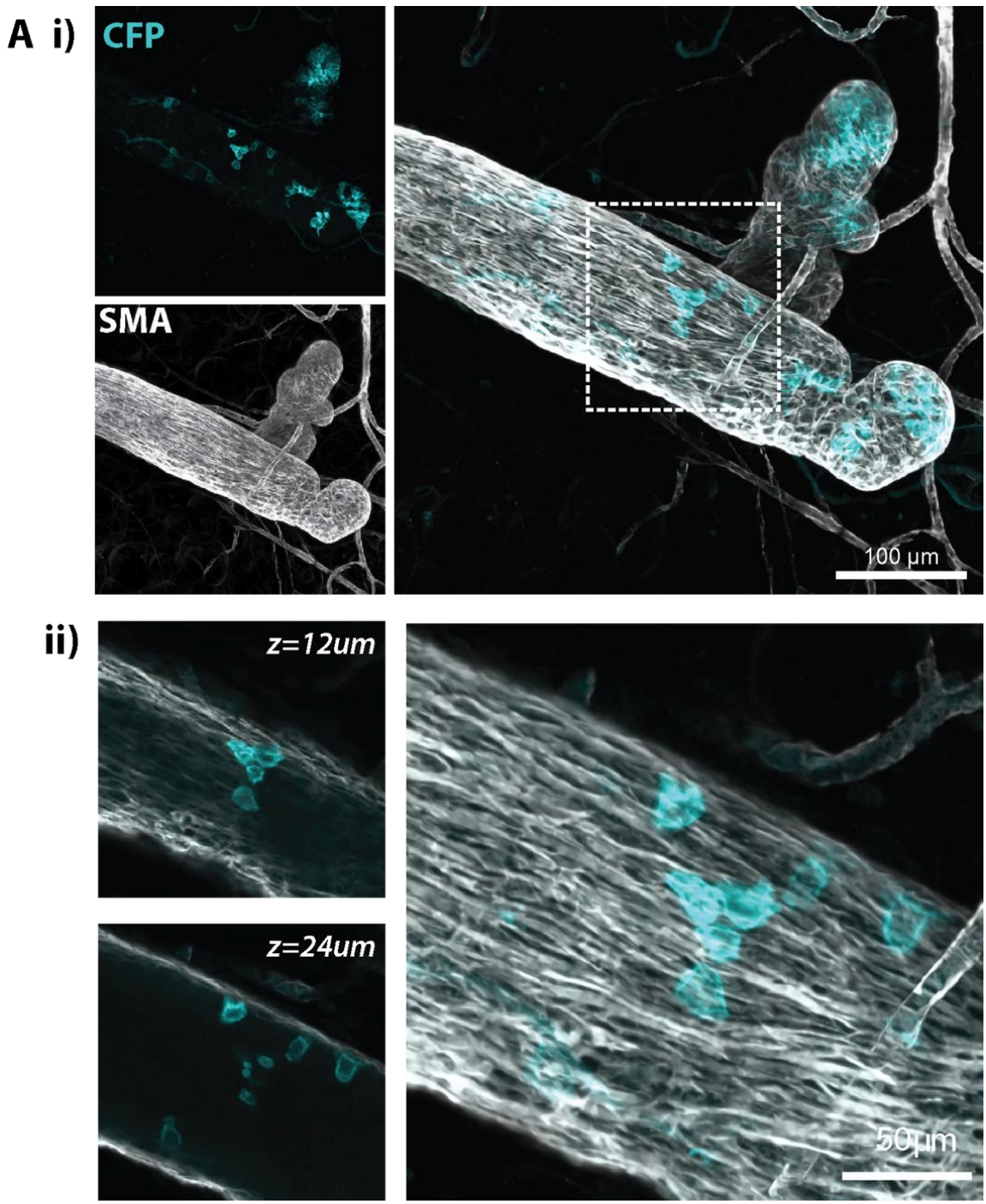


Figure 4.2.2.14. Example of rare CFP labelling in R26R-Confetti mammary glands.

(A) i and ii show maximum-intensity z-projection (single colour and overlay) and optical slices of a demarcated magnified region, inset.

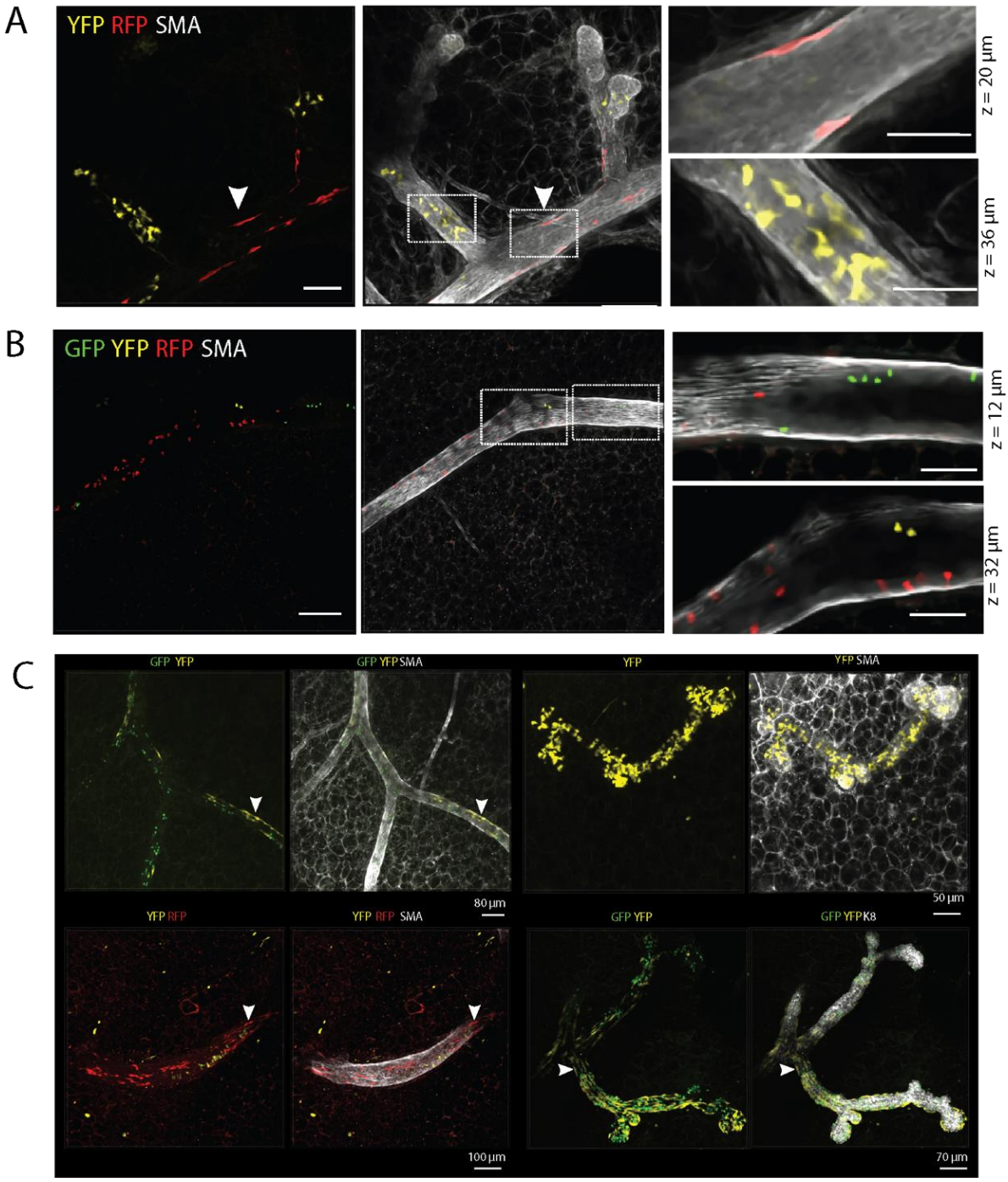


Figure 4.2.2.15. Labelling in R26R-Confetti mammary glands during branching morphogenesis after 1mg tamoxifen induction

(A), **(B)** and **(C)** show representative images of maximum-intensity z-projections and optical slices of higher density labelling following induction of lineage tracing with 1mg of tamoxifen at 4 weeks of age, and imaged 3 weeks later. **(A)** shows a region containing YFP luminal cells and RFP basal cells (arrowhead) populating three branches, interspersed with unlabelled cells. **(B)** shows a region containing GFP, YFP and RFP luminal cells in a single branch. Scale bars, 100µm (overview) and 50µm (inset). Images are representative of three mice.

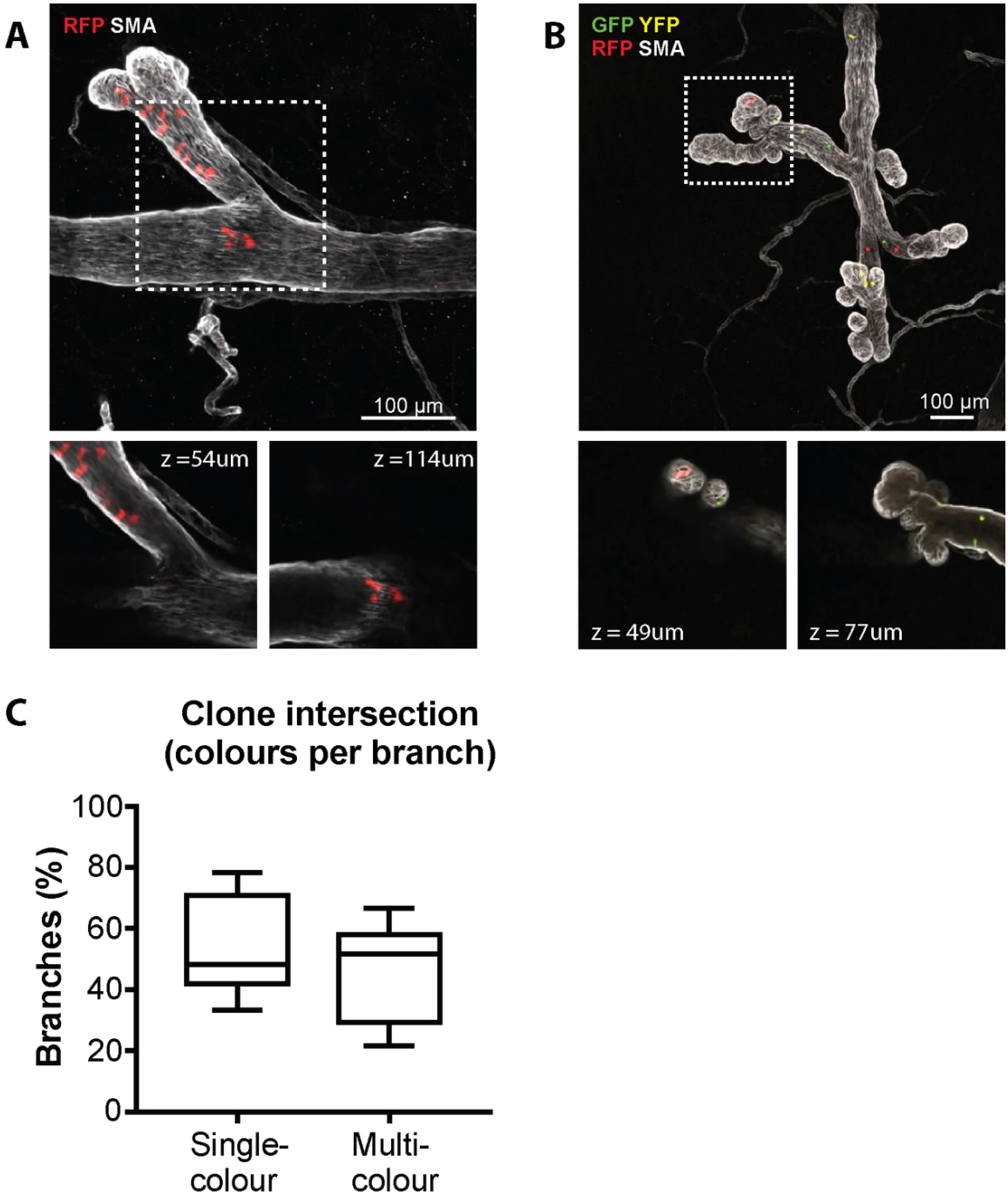


Figure 4.2.2.16. Labelling in R26R-Confetti mammary glands during branching morphogenesis after 0.5mg tamoxifen induction

(A) Example of single-colour branches and multicoloured branches **(B)**. Images show maximum-intensity z-projections and optical slices of a region of interest (boxed and enlarged in panels below). **(C)** Quantification of labelling, showing the percentage of single- and multicolour branches in pubertal mice.

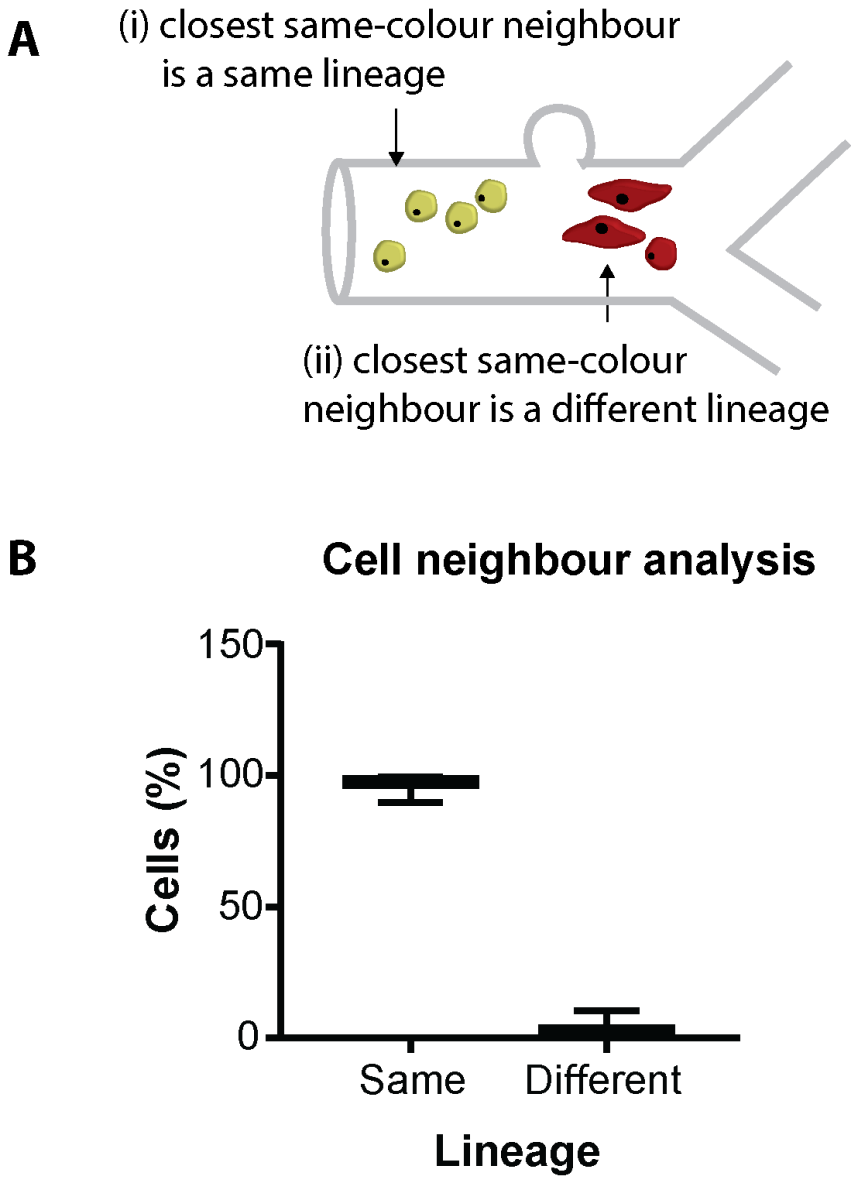


Figure 4.2.2.17. Lineage analysis in R26R-Confetti mammary glands during branching morphogenesis after 0.5mg tamoxifen induction

(A) Schematic outlining method of cell neighbour analysis carried out on R26-Confetti mice. **(B)** Results of nearest cell neighbour analysis, which revealed that the majority of labelled cells had a same colour labelled neighbour of the same lineage. Box extends from the 25th to 75th percentiles and whiskers indicate minimum to maximum values. Data from 1419 cells distributed across 130 branches from randomly selected 3D images (n=5 mice).

4.2.3 Mammary stem cell tracing beyond pubertal morphogenesis

The mammary gland is traditionally thought to maintain a relatively quiescent state when not undergoing the morphogenic events during pubertal or alveolar development. However, the gland in fact undergoes small cyclical changes with oestrous cycling, and as such small changes can occur. Consequently, we wanted to investigate the fate of mammary stem or progenitor cells into adulthood. To do so, we utilised the R26-Confetti model whereby timing can be accurately applied. Lineage tracing was induced in virgin mice at 10 weeks of age, well beyond pubertal development when all TEBs will have regressed, to ensure tracing of stem or progenitor cells responsible for adult homeostasis. The mammary glands were then examined one year later.

Initially, glands were optically cleared using the SeeDB protocol, and co-stained with DAPI and SMA or K8. However, due to the vast increase in size of the mammary glands at one year of age compared with the pubertal mammary gland, and the consequential increase in adipose tissue, imaging depth was severely affected, making lineage identification difficult (Figure 4.2.3.1a). However, this still allowed visualisation of a number of small clones, with some multicolour regions detected (Figure 4.2.3.1b). However, the majority of labelled regions contained single colour clones (Figure 4.2.3.1c,i and ii).

Consequently, we opted to use the CUBIC protocol which provides superior imaging depth due to the use of detergents. Accordingly, this provided better resolution of the adult gland (Figure 4.2.3.2a). Interestingly, despite the better depth achieved, only small clones were detected still. Some of the regions observed were of single colour (Figure 4.2.3.2b), but a number of regions were seen with more than two colours (Figure 4.2.3.2c,d). Through co-localisation, these clones all appeared to be lineage restricted, with mostly luminal clones seen (data from 12 mammary glands from 3 mice).

Overall, this suggests that the mammary stem or progenitor cells involved in homeostasis of the adult gland and oestrous cycling do not have a large capacity. One preliminary hypothesis may be that in fact the turnover of the gland during oestrous cycling means that progeny of MaSCs are constantly being replaced. However, the reduced imaging depth may mean other labelling events may have not been detected. Moreover, more mammary glands would need to be examined, as rare events would not be detected by this method.

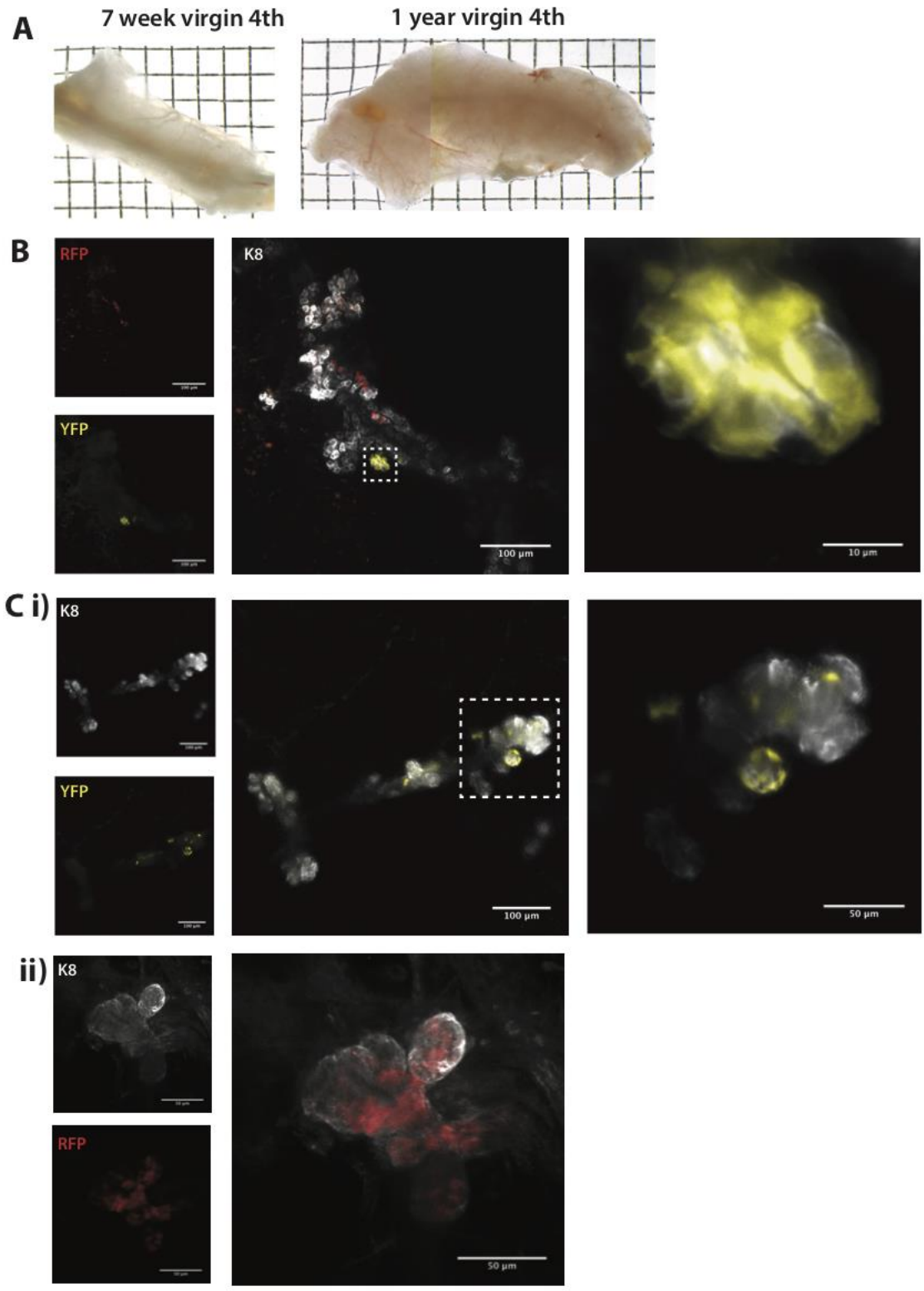


Figure 4.2.3.1 Lineage tracing in R26R-Confetti mammary glands during adult homeostasis, imaged with SeeDB optical clearing

(A) Size difference in inguinal (4th) mammary glands from a 7 week virgin (left) and 1 year virgin (right) **(B)** Representative maximum-intensity z-projections and optical slices of multi-coloured labelled region and **(C)** i and ii regions with labelling of only single colour (YFP and RFP, top and bottom panel respectively)

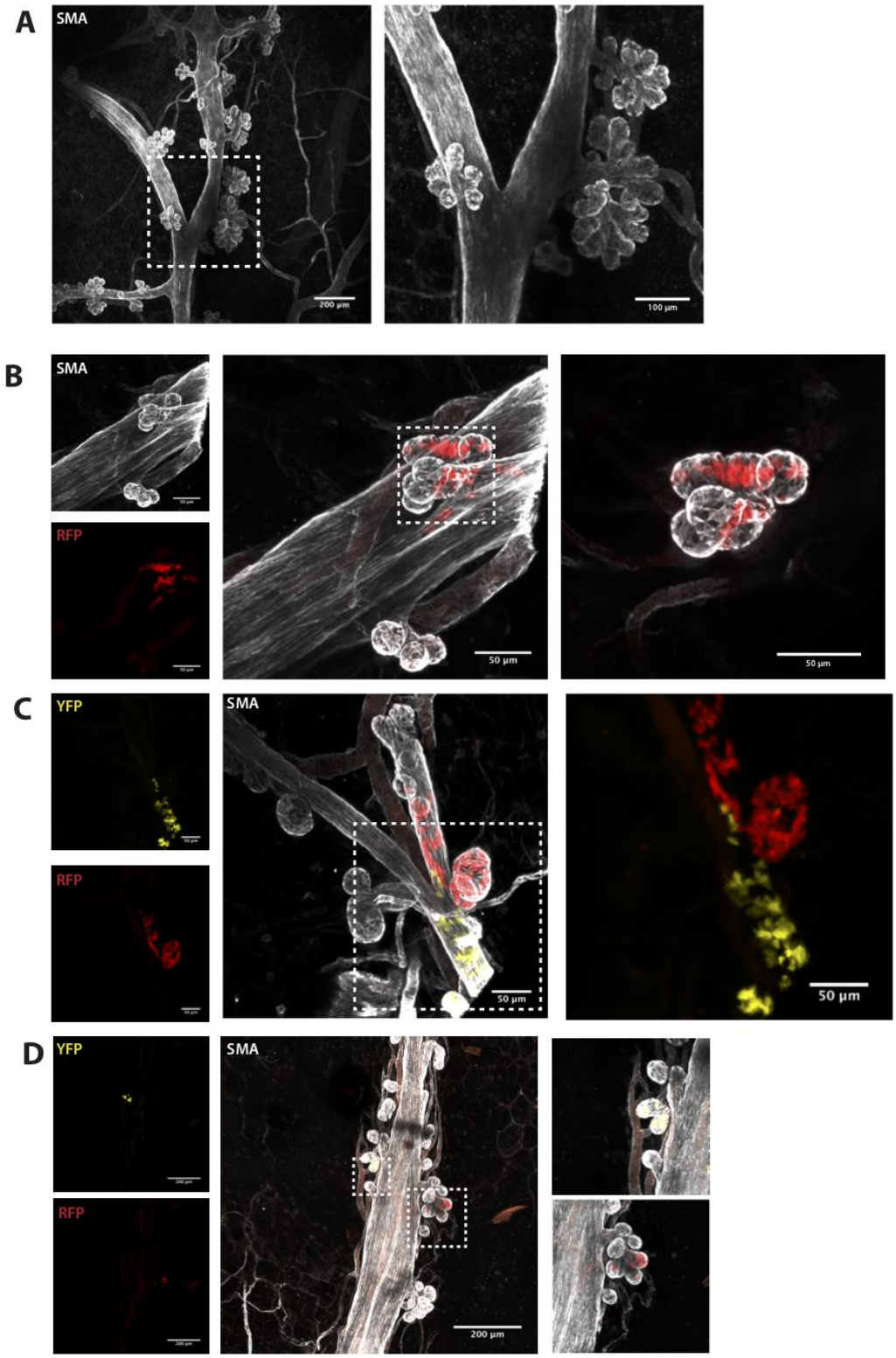


Figure 4.2.3.2 Lineage tracing in R26R-Confetti mammary glands during adult homeostasis, imaged with SeeDB optical clearing

(A) Maximum-intensity z-projection showing an example of superior imaging of 1 year old mammary gland staining with SMA and cleared using CUBIC optical clearing. **(B)** Representative maximum-intensity z-projections and optical slices of single coloured region, with RFP cells seen on side branches. **(C)** and **(D)** Regions with labelling multi-colour labelling. Data from 12 mammary glands from 3 mice.

4.2.4 Tracing alveolar mammary stem cells

The other main morphogenic event of the mammary gland is pregnancy and lactation, whereby the established adult ductal tree undergoes enormous amounts of proliferation and tertiary branching and the milk-producing alveoli are formed at the branch tips, fulfilling the evolutionary purpose of the mammary gland. Like the pubertal mammary gland, the identity of the stem and progenitor cells involved in fulfilling this morphogenesis are still under much debate (Lloyd-Lewis et al., 2017). As such we opted to use the neutral lineage tracing approaches, combined with 3D imaging, to characterise the MaSC and/or progenitors involved in this vital process.

4.2.4.1 R26$^{[CA]30}$ YFP

The first model utilised to study mammary alveologenesis was the R26$^{[CA]30}$ YFP model. Mammary glands were harvested between lactation days 2 and 4 and SeeDB optical clearing utilised for 3D imaging with confocal microscopy. Labelling was observed throughout the lactating gland, with both luminal and basal clones detected (Figure 4.2.4.1a, left and right panels respectively). Variable sizes of clones were observed (Figure 4.2.4.1b), with the smaller ones reflecting either more recent slippages or stem or progenitors with limited potential.

Of note, slippage was observed to be more frequent in the lactating glands when compared to the pubertal mammary gland – which is not unexpected given the increased number of cells and extensive proliferation during gestation to produce the expansion of the ductal tree and alveoli. A pattern of stochastic labelling was again observed, in a similar manner to the pubertal gland, whereby labelled cells were intermixed with the progeny of unlabelled stem or progenitor cells. This suggests that multiple stem or progenitor cells are also involved in alveologenesis of the mammary gland. Moreover, YFP+ cells were distributed with differential labelling patterns: the majority of alveoli were partially YFP+ (Figure 4.2.4.2a,i), but alveoli that were where almost entirely labelled were also seen (Figure 4.2.4.2a,ii). Interestingly, a third labelling pattern was detected, with only a single YFP+ cell per alveolus (Figure 4.2.4.2a,iii).

In total, 4 large clonal regions were detected from the 80 mammary glands examined (n=10 mice). Although low, this highlights the rare nature of the slippage and as such ensures confidence in single cell labelling. It may also suggest that MaSCs with such a large capacity are infrequent, with the majority of cells contributing to smaller portions of the alveolar units. The

capacity of these MaSCs were evaluated by quantifying the area of the gland that the clones occupied, and the potency by assessing the lineage of the cells (summarised in Figure 4.2.4.2b).

75% of the large clonal regions detected contained entirely cells of the luminal lineage. In one particularly vast clonal region, the progeny of the lineage restricted MaSCs comprised up to more than 169 individual alveolar units (Figure 4.2.4.3a). The other two luminal clones were slightly smaller, occupying more than 68 individual alveolar units (Figure 4.2.4.3b and c). Interestingly, the differential expression of K8 in luminal cells, as was seen in pubertal ducts, was also seen in lactational alveoli (Figure 4.2.4.4a). However, rather than being present in similar proportions as in the pubertal gland, the majority of luminal cells were $K8^{lo}$, with only single rare $K8^{hi}$ cells present on each alveolus. Given these $K8^{hi}$ cells also co-stained with PR, they may comprise a similar hormone sensing lineage of luminal cells as was seen in pubertal ducts (Figure 4.2.4.4a, arrow heads). Furthermore, luminal YFP clones comprised both of these populations of K8+ luminal cells (Figure 4.2.4.4b), suggesting that lineage restricted luminal MaSCs contribute to both the $K8^{lo}$ and $K8^{hi}$ luminal subtypes.

Interestingly, out of the 80 mammary glands examined only one large basal clone was seen, identified by co-staining with SMA. Similar to the luminal clones seen, the progeny of this lineage restricted MaSC occupied a vast number of alveolar units, comprising over 100 individual alveoli, some of which were entirely labelled by YFP+ cells (Figure 4.2.5.5i and ii). In order to assess the distribution of clonal patterns further, the number of alveoli that were fully populated by YFP+ cells of a single lineage (full) or populated by both YFP+ and unlabelled cells of a single lineage was calculated in the four large clones (Figure 4.2.4.5b). This demonstrated that the majority of alveoli were only partially labelled - in both luminal and basal clones – suggesting that most alveoli are polyclonal, comprised of the progeny of multiple lineage restricted MaSCs.

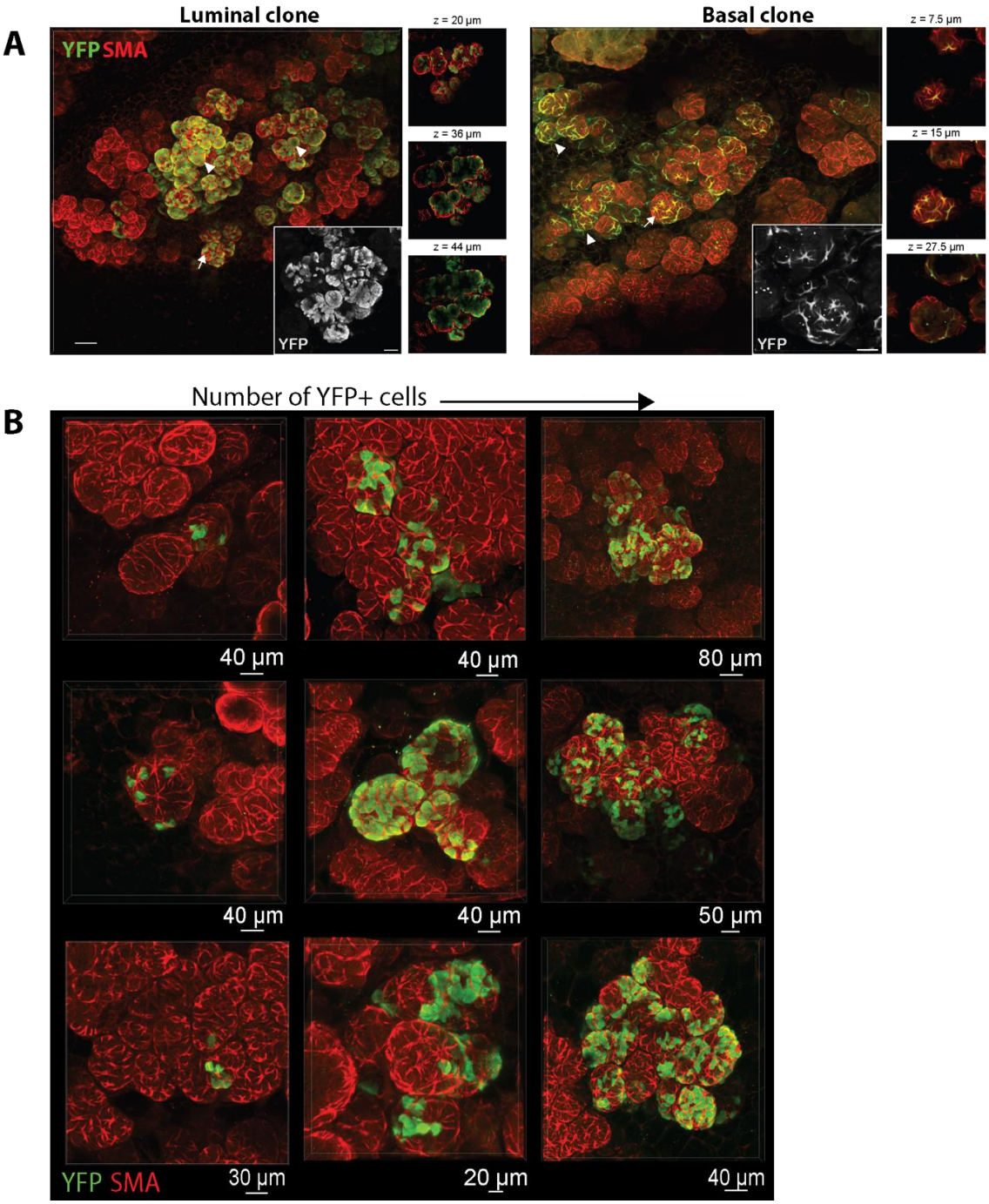
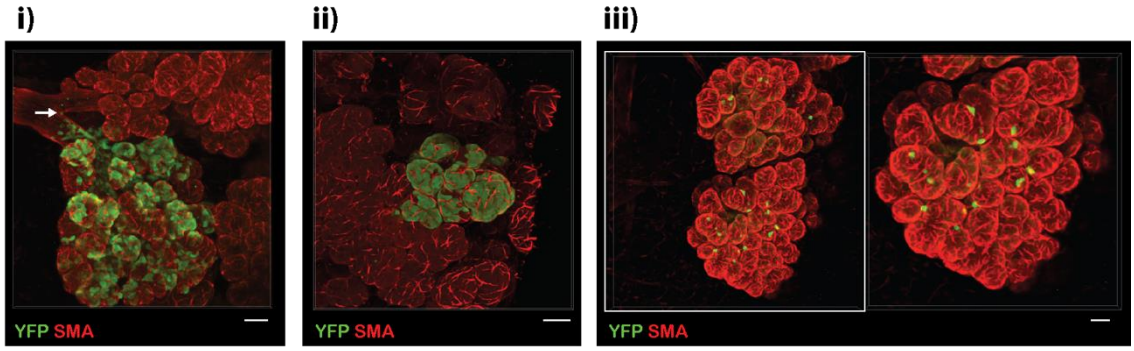


Figure 4.2.4.1 Overview of labelling in the R26$^{[CA]30}$ YFP mouse model for lineage tracing during lactation

(A) Representative maximum-intensity z-projection images, revealing extensive labelling in the lobuloalveolar network of a luminal (left) and basal (right) clonal region in independent lactating mammary glands. Labelled alveoli were populated by both lineage-restricted YFP and unlabelled cells (arrow), with occasional alveoli observed that were fully populated by YFP cells of a single lineage (arrowhead). Scale bar, 100 mm (overview) and 40 mm (inset). **(B)** Representative maximum-intensity z-projection, showing examples of labelled regions of increasing sizes observed.

A



B

R26[CA]30EYFP				
Total no.	10 mice (80 mammary glands)			
Linear length (mm)	> 2	> 1	> 1	> 2
No. alveoli	> 169	> 68	> 69	> 100
Composition	Lu	Lu	Lu	Ba

Figure 4.2.4.2 Examples of different labelling patterns in the R26[CA]30 YFP mouse model for during lactation

(A) Representative maximum-intensity z-projection showing the different labelling patterns: i) shows uneven contribution of a single labelled ductal cell to different lobuloalveolar structures; arrow indicates the presumptive YFP cell of origin at the ductal branch point. (ii) shows a rare instance of progeny from a single luminal YFP cell contributing almost entirely to the luminal lineage of 2–4 alveoli within a single lobule and (iii) shows an example of a single labelled YFP luminal cell that contributed one luminal daughter cell to multiple alveoli in a lobule. Scale bars, 100µm. **(B)** Tabulated summary of the four large clonally marked regions observed from the analysis of 80 mammary glands from 10 R26[CA]30 YFP mice during lactation.

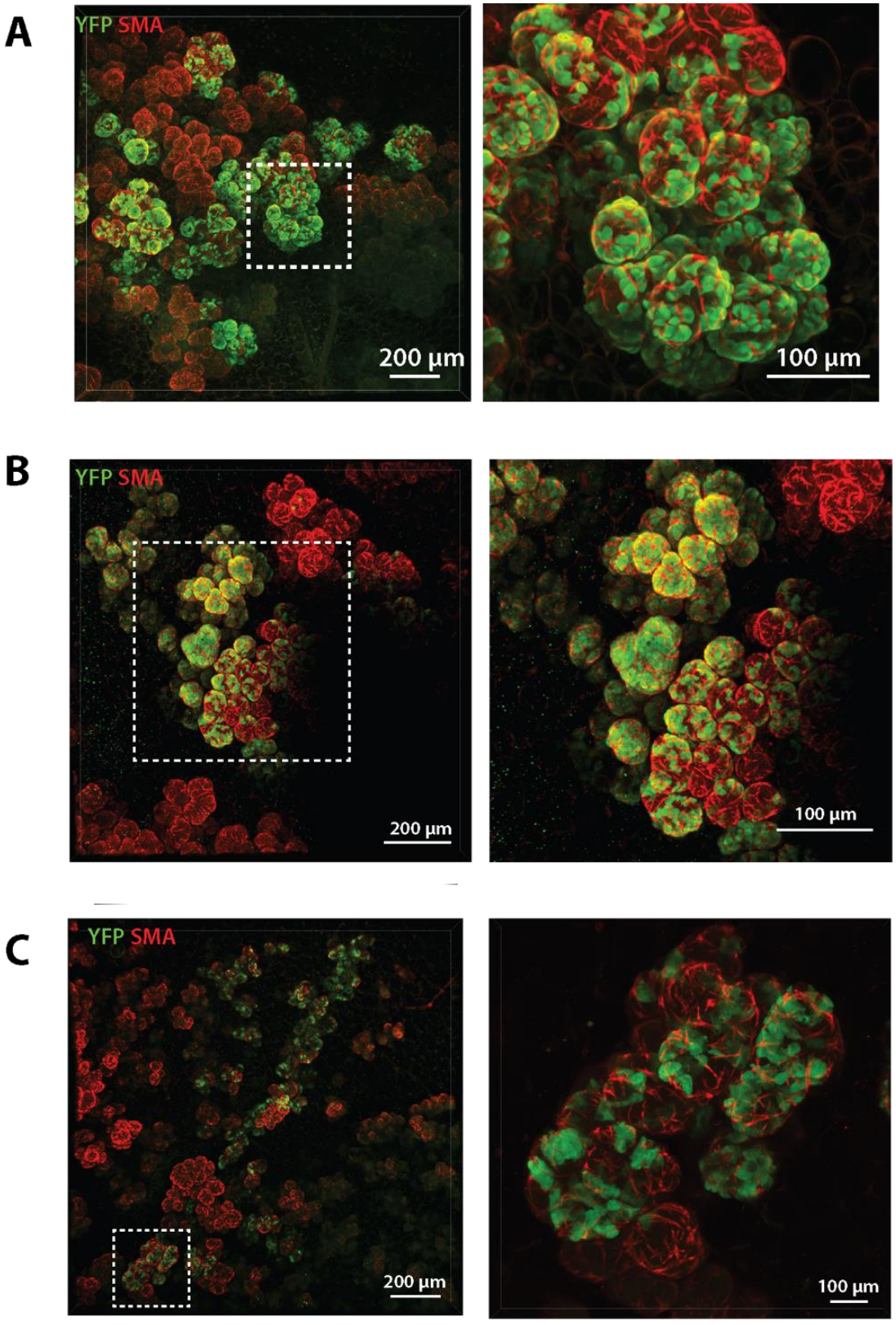


Figure 4.2.4.3 Luminal clonal regions in the R26$^{[CA]30}$ YFP mouse model for during lactation

(A), **(B)** and **(C)** Maximum-intensity z-projections showing the extent of the three independent luminal clonal regions detected from analysis of 80 mammary glands from 10 R26$^{[CA]30}$ YFP mice during lactation.

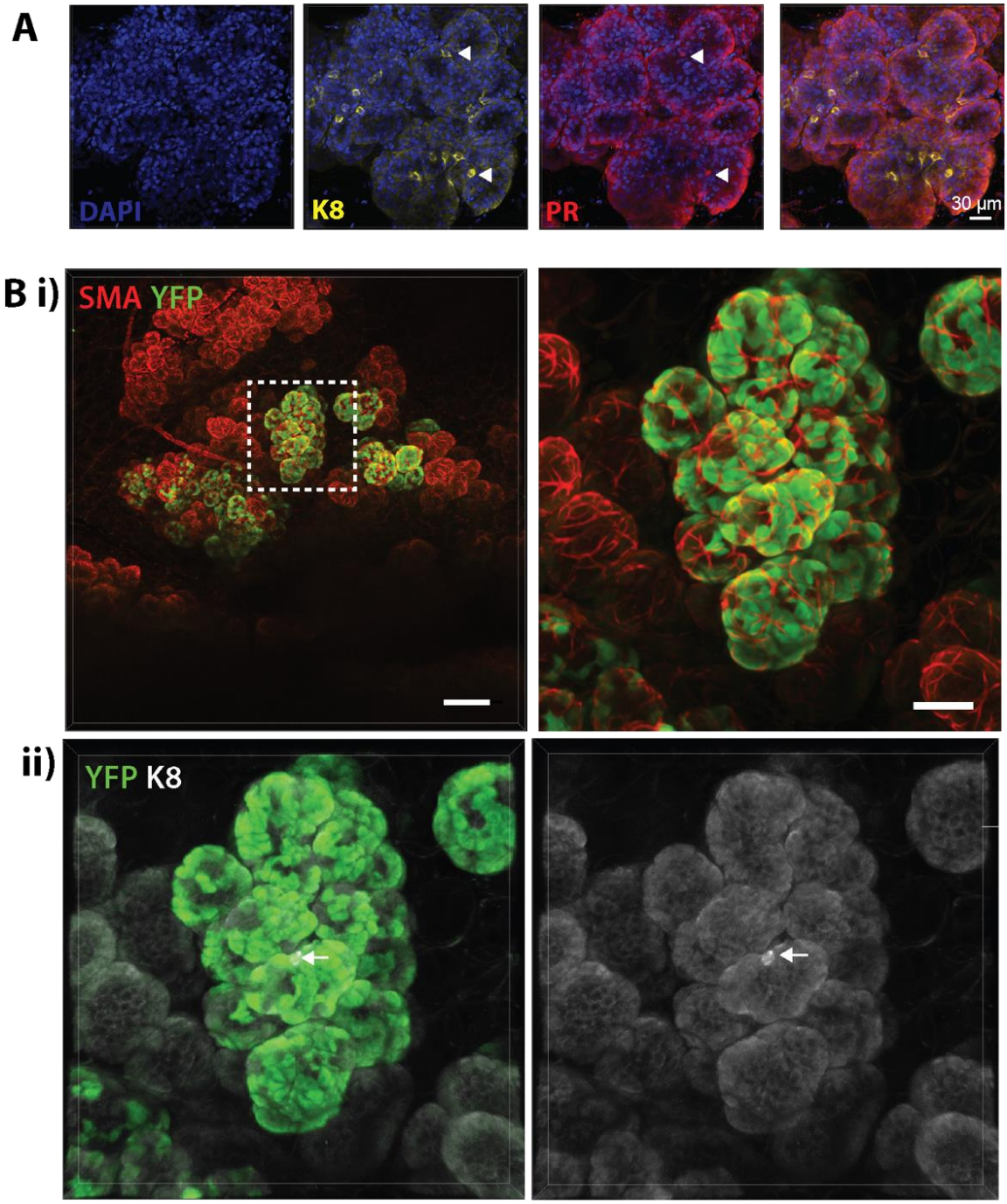


Figure 4.2.4.4 Differential K8 expression seen in luminal cells of lactating R26$^{[CA]30}$ YFP mice

(A) Maximum-intensity z-projections showing Rare K8hi / PR+ luminal alveolar cells are observed in lactating mammary tissue. **(B)** Maximum-intensity z-projections showing that lineage restricted luminal MaSCs can contribute to both K8hi and K8lo populations. i) shows the YFP labelled region SMA co-staining with magnified view, and ii) shows K8 co-staining, with K8hi cells (arrows). Scale bars, 150µm and 50µm (inset).

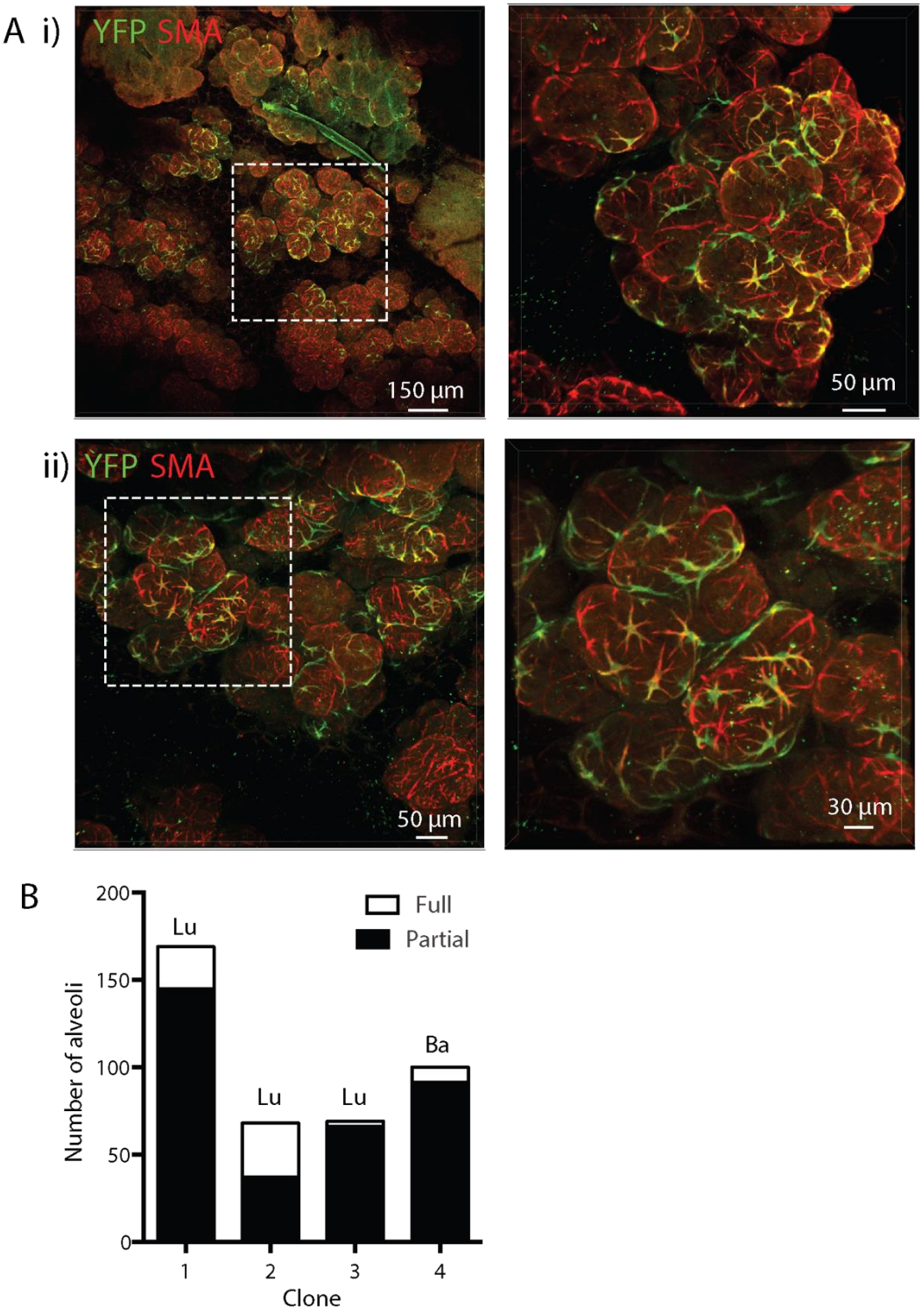


Figure 4.2.4.5 Basal clonal regions in the R26$^{[CA]30}$ YFP mouse model for during lactation

(A) i and ii show maximum-intensity z-projections showing the extent the single basal clonal region detected from analysis of 80 mammary glands from 10 R26$^{[CA]30}$ YFP mice during lactation. **(B)** Quantification of the number of alveoli that were fully populated by YFP cells of a single lineage (full) or populated by both labelled and unlabelled cells of a single lineage (partial). Lu, luminal; Ba, basal.

4.2.4.2 R26-Confetti lineage tracing in the lactational mammary gland

To confirm the findings seen with the $R26^{[CA]30}$ YFP model and to ensure accurate timing we next employed the R26-confetti model to study the role of adult MaSCs in alveologenesis. Adult virgin $R26CreER^{T2}$;R26-Confetti mice (approximately 12 weeks of age) were injected with low-dose tamoxifen (1 mg per mouse, equivalent to roughly 40-50 μg/g body weight) and mated two weeks later. Following littering, mammary glands were harvested during lactation days 4 to 5 (Figure 4.2.4.6a). Importantly, no labelling was observed in glands taken from vehicle injected control mice (Figure 4.2.4.6b,I and ii). Reporter expression in the mammary glands was observed after a short chase period of 2 days to examine initial labelling events (Figure 4.2.4.7). Small regions of mostly individual cells were identified.

Glands taken during early lactation were then examined for labelling events. Similar to the pattern seen in the $R26^{[CA]30}$ YFP model, large clonal regions were observed, occupying multiple alveolar units. Some multicolour alveoli were observed (Figure 4.2.4.8a), however the majority of alveoli contained cells labelled with a single colour (Figure 4.2.4.8b,c). This was confirmed by subsequent quantification of alveoli, with on average 96.6% of alveoli containing single-coloured labelling, providing confidence in the clonal labelling method (Figure 4.2.4.9a,i).

In order to analyse potential potency of the mammary stem or progenitor cells labelled we next analysed the lineage of these single colour labelled alveoli; strikingly, this revealed that only 0.1% contained both luminal and basal cells of the same colour (Figure 4.2.4.9a,ii). Together, these data support the data from the $R26^{[CA]30}$ YFP data, in that the MaSCs involved in forming the milk-producing alveoli in lactation are also lineage restricted, similar to the morphogenesis events in embryogenesis and pubertal mammary gland development.

The final aspect we sought to investigate was the clonality of singular alveolar units, as was shown in the $R26^{[CA]30}$ YFP model. Therefore, the proportion of full versus partially labelled alveoli was quantified. Similar to the $R26^{[CA]30}$ YFP model, this showed, in both basal and luminal clones, that the vast majority of alveoli were partially labelled. Only 2.3% of luminal and 0.5% of basal clones occupied entire alveolar units, out of 793 and 173 units examined, respectively (Figure 4.2.4.9b, i and ii).

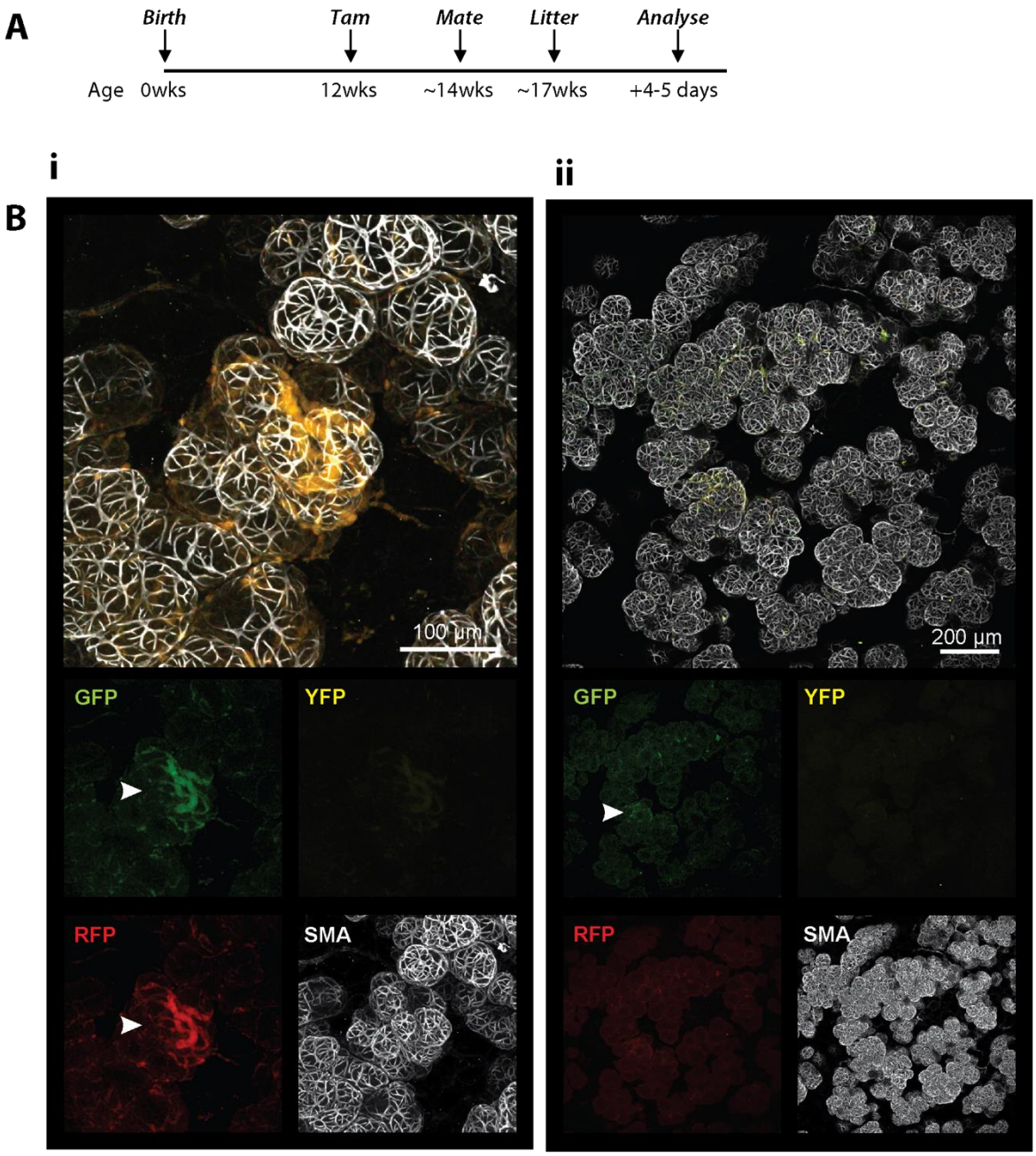


Figure 4.2.4.6 Overview of lineage tracing in R26R-Confetti mammary glands during lactation

(A) Schematic of lineage tracing schedule: 1mg of tamoxifen was administered in adult virgins, at approximately 12 weeks of age, and mice were mated 2 weeks later (at approximately 14 weeks of age). Mammary gland tissue was harvested during lactation days 4 – 5. **(B)** i and ii) Images showing maximum intensity z-projections, demonstrating the absence of labelling in control mice injected with oil (vehicle). Arrowheads show non-specific background fluorescence.

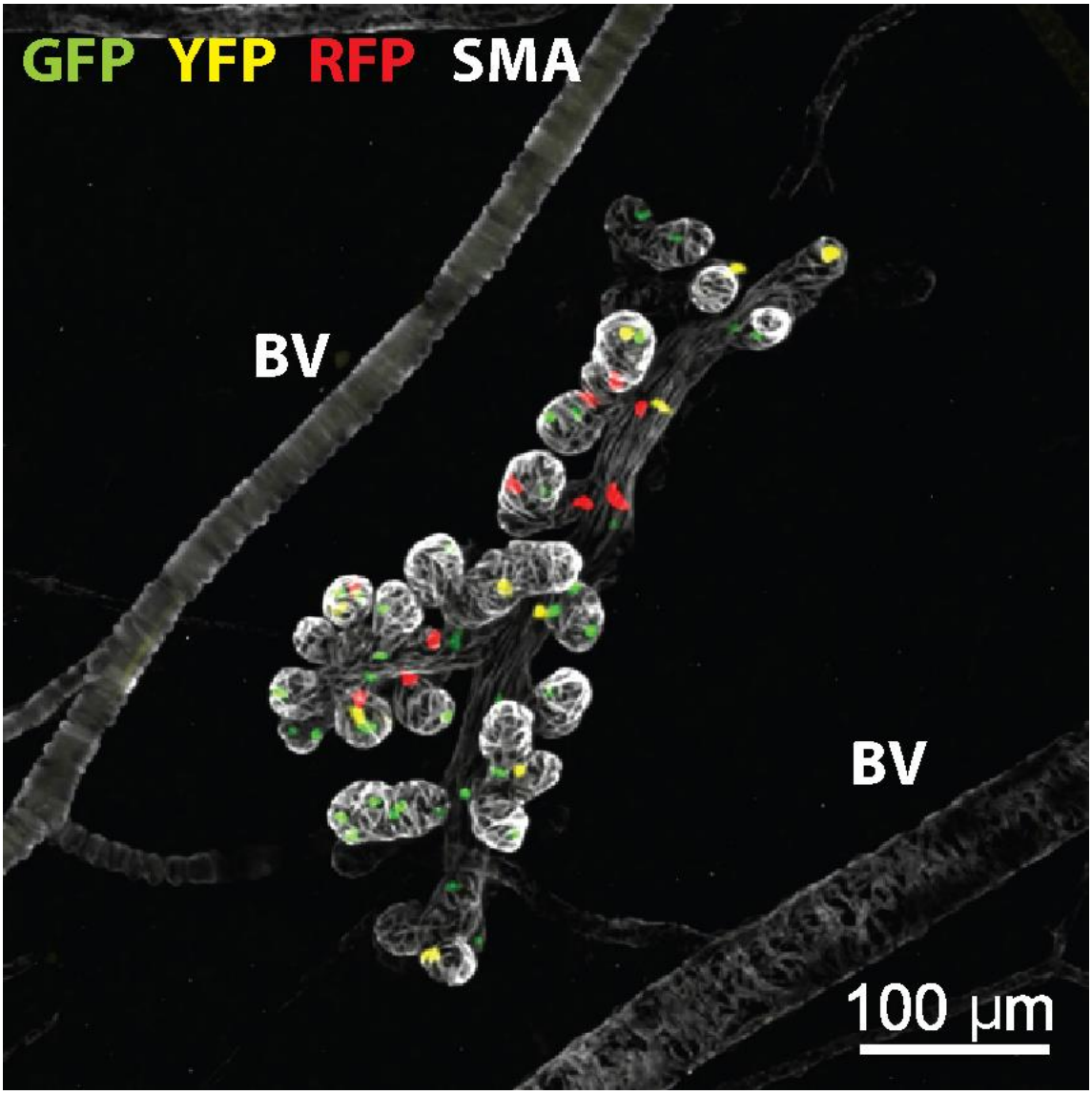


Figure 4.2.4.7 Lineage tracing in R26R-Confetti mammary glands during lactation

Maximum intensity z-projection images of mammary glands from approximately 12 week old R26-Confetti mammary glands, showing the level of labelling following a short (2 day) chase. (BV=blood vessels).

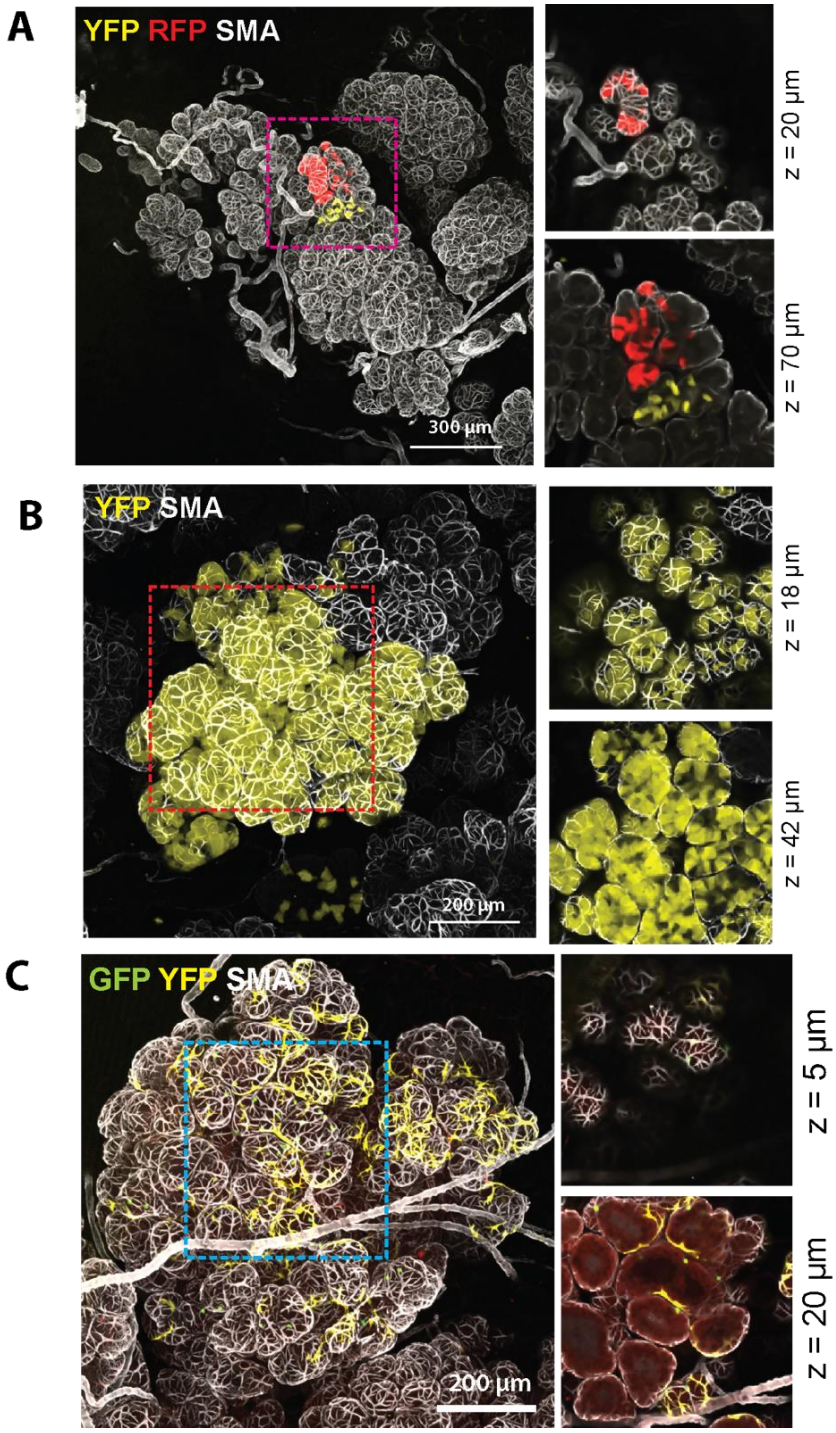


Figure 4.2.4.8 Example of clonal regions seen in lineage tracing of R26R-Confetti mammary glands during lactation

(A) Representative maximum intensity z-projection, showing single-colour luminal YFP+ and RFP+ adjacent alveoli within a larger lobuloalveolar structure. **(B)** Representative maximum intensity z-projection, showing single coloured alveoli, with YFP+ cells throughout a lobuloalveolar unit. **(C)** Example of a multi-colour basal GFP+ and YFP+ alveoli. Images show maximum- intensity z-projections and optical slices of a region of interest (boxed and enlarged in the right-hand panels).

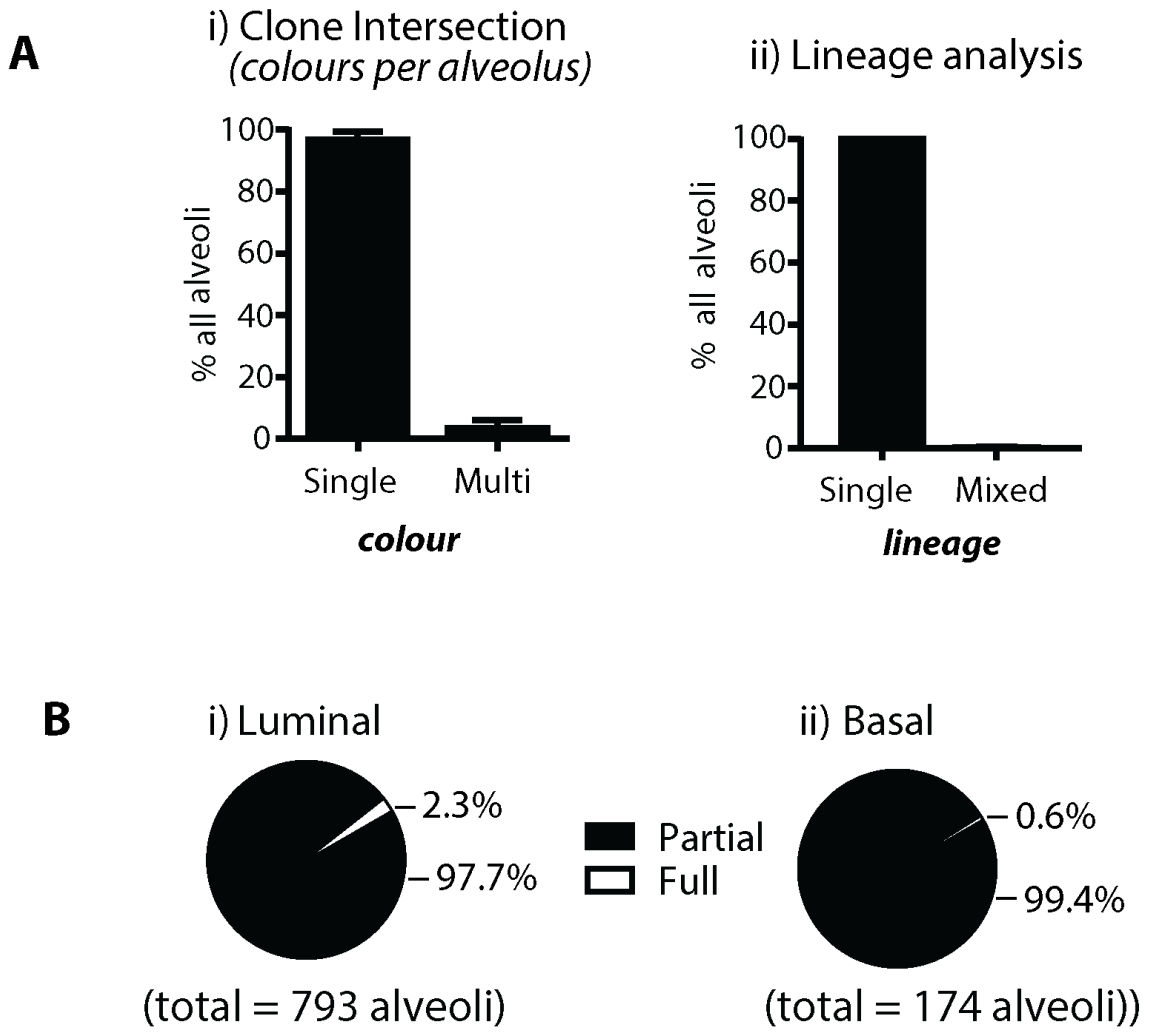


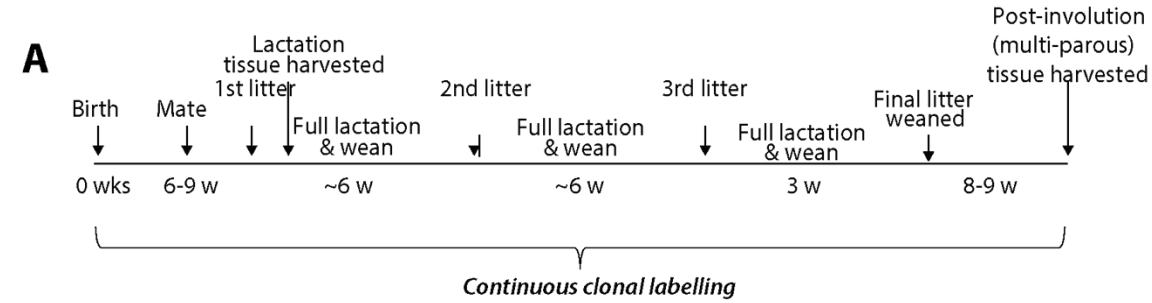
Figure 4.2.4.9 Quantification of lineage tracing in R26R-Confetti mammary glands during lactation

(A) Quantification of R26R-Confetti labelling during lactation: **i)** graph showing the percentage of single- and multi-coloured alveoli and **ii)** graph showing the percentage of labelled alveoli in which the same colour cells were the same lineage (i.e. all luminal or all basal), or where same-colour cells were mixed lineage (both luminal and basal). Data are mean±s.e.m. **(B)** Fraction of alveoli that were fully populated by single-colour labelled cells of a single lineage (full) versus those populated by both single- or multi-coloured labelled cells and/or unlabelled cells of a single lineage (partial) in both luminal (i) and basal (ii) clones. Data represent 1016 alveoli from randomly selected 3D images (n=3 mice).

4.2.4.3 R26$^{[CA]30}$ SYNbglA

Finally, we utilised the R26$^{[CA]30}$ SYNbglA model to analyse MaSC presence after multiple pregnancies. For this, adult virgin mice that were hemizygous for R26$^{[CA]30}$ SYNbglA were mated and underwent full lactation for an average of 3 weeks, followed by natural weaning. Mating was repeated a further two times (Figure 4.2.4.10a), resulting in R26$^{[CA]30}$ SYNbgl mice that had undergone three full gestation, lactation and involution cycles. This allowed us to examine the presence of clonal regions that were still present following multiple involutions.

We first identified labelling during the first lactation to validate the use of this model, with multiple, small regions seen (Figure 4.2.4.10b). Following this, we observed staining in the multiparous mice. Despite the multiple pregnancies and consequently multiple expansions of the gland, only small labelled regions were observed (Figure 4.2.4.11). This is potentially due to remodelling of the gland involved in the process of involution. However, this may also support the notion that some cells remain following multiple involutions. While these may represent the progeny of a long-lived MaSC, since this is a continuous clonal labelling model, the originating cell may have been labelled at any stage of the study (which totalled approximately 8 months).



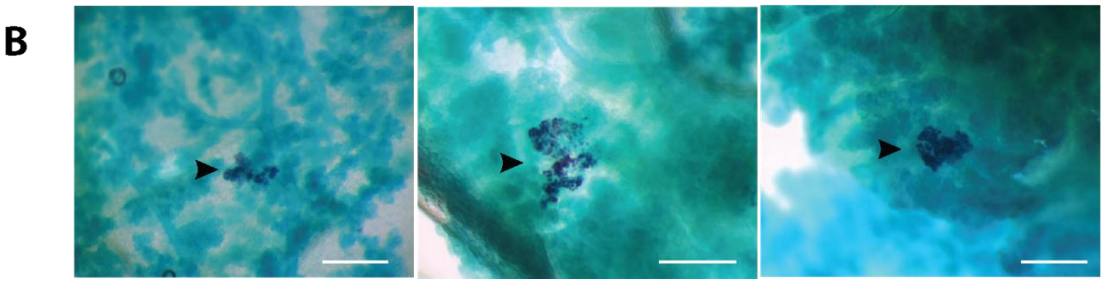


Figure 4.2.4.10. Use of R26$^{[CA]30}$ SYNbglA mice to study MaSCs involved in multiple gestation / lactation / involution cycles

(A) Schematic showing the experimental protocol for obtaining lactating multiparous mammary tissue: mice of reproductive age were mated and lactating tissue harvested from some mice on lactation day 4. A cohort of mice were taken through three pregnancies each followed by a full natural wean (approximately 20 days). Mammary tissue was harvested 8-9 weeks after the final wean. n=3 mice for each time-point. **(B)** Representative images showing initial SYNbglA labelling after 4 days of first lactation. Scale bar shows 500 µm.

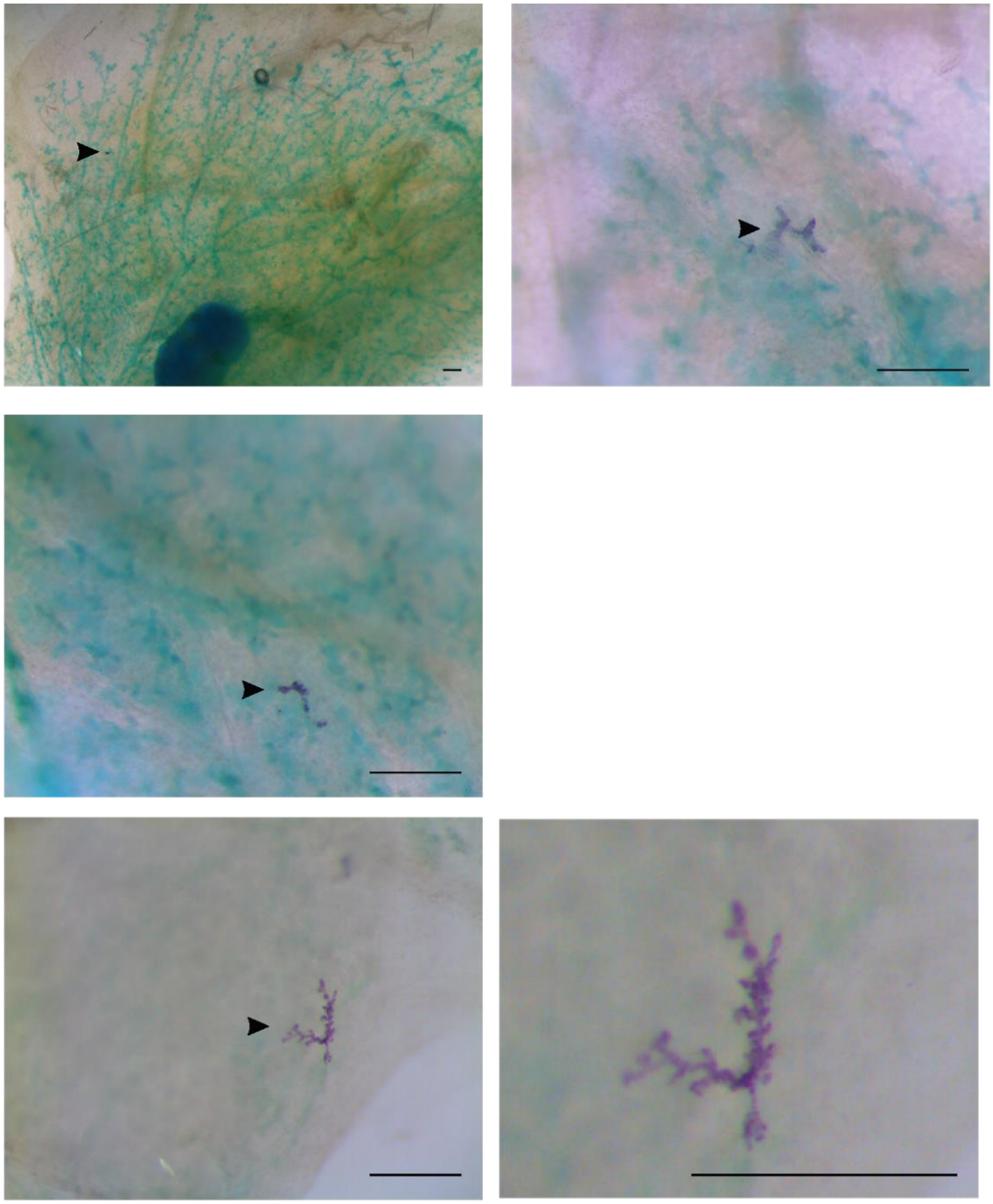


Figure 4.2.4.11 Representative labelling in R26$^{[CA]30}$ SYNbglA mice after involved in multiple gestation / lactation / involution cycles

Representative stereomicroscopy images of mammary glands following three full gestation / lactation / involution cycles. These SYNbglA cells (purple, arrowhead) may represent the progeny of a long-lived MaSC – however since this is a continuous clonal labelling model, the originating cell may have been labelled at any stage of the study.

4.3 Discussion

4.3.1 Summary

The identity of MaSCs has been a long debated topic since their existence was first demonstrated nearly 60 years ago (DeOme et al., 1959). Consequently, despite research into the subject having spanned a number of decades by a number of groups, considerable uncertainties and disparities between results remain. Recently, the molecular identity and potency of MaSCs in particular has come under scrutiny. Prime amongst the debate is the fate of adult MaSCs; a number of groups have shown data to support either side of the potency argument, demonstrating MaSCs being either unipotent or bipotent (Van Keymeulen et al., 2011; van Amerongen et al., 2012; Rios et al., 2014; Wang et al., 2015; Wuidart et al., 2016; Scheele et al., 2017).

One proposed reason for these inconsistencies is due to the experimental approaches used; previous lineage tracing has relied on presumed lineage-restricted promoters to drive reporter gene expression in a significant proportion of MaSCs for population-based fate-tracking. In using this approach, the probability of two or more clones arising in the same region is high, leading to the potential convergence of clones which can result in incorrect interpretation of the data. In this case, labelling and subsequent expansion of two single lineage-restricted MaSC could resemble clonal expansion of a single bipotent MaSC.

Furthermore, recent work has highlighted that various markers presumed to be lineage restricted may not be so specific (Koren et al., 2015; Wuidart et al., 2016). Indeed, this chapter has demonstrated differential expression of K8, whereby differential expression is seen in luminal cells throughout the pubertal and adult gland. Consequently, use of K8 would lead to disparate labelling of luminal cell populations in fate mapping studies.

Additionally, recent work from our laboratory has also highlighted that K14, a marker of basal cells, can be detected in rare luminal cells within the rabbit mammary epithelium (Hughes and Watson, 2018). Although murine and rabbit mammary glands have slight differences in embryonic development and ductal morphology (Propper et al., 2013), they share a number of similarities, such as marker expression. Consequently, this would need to be investigated further in the murine gland. However K14 expression in luminal cells has been previously demonstrated in the pre-pubertal mouse mammary gland (Sun et al., 2010; Boras-Granic et

al., 2014). Accordingly, use of these as promoters to drive tracing in mammary cells may in fact not result in entirely specific recombination and thus lineage tracing with these promoters is not definitive for the assessment of potency.

4.3.2 Overview of main findings

Accordingly, this means that investigating the role of mammary stem and progenitor cells requires the use of novel methods, and multiple models to validate conclusions. Here, we have developed techniques to support the use of neutral lineage tracing in both the $R26^{[CA]30}$ and R26R-Confetti models in an attempt to identify the nature of these cells in the mammary gland without pre-determined bias. Moreover, the use of 3D imaging has provided an unprecedented view of mammary stem and progenitor cell progeny.

Using these methods, we have investigated the main aims of this chapter - namely to characterise the fate of mammary stem and progenitor cells. However, the use these novel approaches outlined here have also resulted in a number of other unanticipated observations and valuable insights. The first, as discussed above, is the differential expression of K8 in luminal cells, with two populations of $K8^{hi}$ and $K8^{lo}$ being identified. Interestingly, these cells are found in similar proportions within virgin ducts, but during lactation the majority of cells in the alveoli express very low levels of K8, and $K8^{hi}$ cells were extremely sparse. Moreover, both populations were seen to be produced from a single luminal stem cell.

Previous work has demonstrated an association between high K8/18 expression and a functionally distinct hormone sensing population of luminal cells expressing $CD24^{hi}$/prominin-1^{+}/$Sca1^{+}$ (Sleeman et al., 2007). Moreover, lineage tracing studies using K8-$CreER^{T2}$ mice have demonstrated preferential genetic labelling of $CD24^{hi}$/$Sca1^{+}$ luminal cells (Koren et al., 2015). Indeed, this raises a number of questions: are the $K8^{lo}$ cells seen in lactational alveoli new cells that arise with the proliferation of the gland during gestation, or are they in fact ductal cells that downregulate K8 expression? If so, at what point during gestation or lactation does this process occur, and how are they affected in the remodelling during involution?

Another noteworthy observation is the intra-ductal stochastic distribution of labelled cells, hypothesised to be a result of the admixing of cells with progeny from an unlabelled stem cell. This demonstrates the presence of multiple stem or progenitor cells within a duct, exhibiting the polyclonal nature of ducts (at least during the timepoints examined). This differs from

other epithelial structures such as the intestinal crypts, which overtime drift to monoclonality (Snippert et al., 2010; Kozar et al., 2013; Nicholson et al., 2018), or the pancreatic ductal tree, whereby the majority of branches are formed from a single stem cell (Sznurkowska et al., 2018).

Conversely, another reason for this pattern may be due to the movement of cells after being laid down during ductal expansion. Indeed, this has been demonstrated using intravital microscopy, during which cells were seen to be highly dynamic and migratory within TEBs, although not within the ducts (Scheele et al., 2017). However, TEBs are especially large structures and it may be that imaging deeper into the fat pad is more challenging and as such more difficult to image the ductal network and detect labelled cells.

Likewise, the inter-ductal dissemination of labelled cells is of interest. Tracing of a single cell during pubertal development demonstrated the random distribution of progeny to multiple ducts. Interestingly, some regions contained a number of labelled cells, while other branches had a much lower density. As well as providing further evidence for the presence of multiple lineage-restricted stem cells, this also leads to intriguing insights into morphogenesis of epithelial organs. Likewise, it raises questions surrounding the differential proliferative and competitive behaviours of stem and progenitor cells, which may be responsible for determining this interesting pattern.

Using 3D imaging algorithms, we performed a volumetric analysis on clonal regions identified to estimate the possible number of mammary stem and progenitor cells. Assuming that all cells have a similar capacity to contribute to the ductal tree, this estimated that at least 20 luminal and 15 basal mammary stem and progenitors contribute to the growth of a major duct. By extrapolation for the entire ductal tree, this would mean that there may be as many as a few hundred unipotent luminal and basal MaSCs/progenitors in the entire gland, which are responsible for driving ductal morphogenesis during puberty.

The exact reason for this multiplicity is still unclear, but given the evolutionary importance of the mammary gland, the presence of multiple MaSCs cells would provide a level of redundancy to guarantee complete development of the ductal tree in puberty and subsequent alveologenesis. This would ensure the production of milk and consequent nourishment of the young. Moreover, it may provide a safeguarding mechanism to protect against oncogenic

transformation. In epithelial tissues which contain fewer stem cells, the effect of oncogenic hits on these rarer cells is much more detrimental (Kozar et al., 2013).

In utilising these agnostic approaches, we have aimed to avoid any aforementioned issues that may have caused previous disparities. However, no model is without caveats and as such it is important to recognise the limitations of the systems used. The $R26^{[CA]30}$ system utilises a continuous clonal labelling approach which allows for random labelling during DNA replication in proliferating cells. Although this model is neutral, in that it utilises the ubiquitous R26 promoter, the inherent nature of the labelling system means that it is biased towards proliferative stem or progenitor cells. Consequently, any MaSC which is quiescent will not be labelled with this model. Indeed, this feature may explain the fact that less basal cell clones were detected, given the reduced proliferation that has been shown compared to the luminal compartment (Giraddi et al., 2015; Shehata et al., 2018). Therefore, this approach would not label possible bipotent embryonic MaSCs that may persist after birth and remain quiescent during the time periods examined (Boras-Granic et al., 2014).

Another restriction of the $R26^{[CA]30}$ model is the inability to time tracing events. Accordingly, the R26R-Confetti model was used to supplement this and allow tracking of cells within a known time period. This requires the use of tamoxifen for Cre-mediated recombination and subsequent expression of the confetti reporter, which can be problematic when studying mammary gland development (Rios et al., 2014). The use of high doses of tamoxifen (above 5mg / per 25mg body weight) has been shown to increase proliferation in the short term, and caspase-3-mediated apoptotic cell death in the medium term (Shehata et al., 2014). However, this was shown to not affect pubertal development in the long run. Moreover, low, single doses of tamoxifen were utilised here, which was shown not to affect ductal or alveolar development in wholemounts. However, the development of more doxycycline-induced Cre expression mouse models may reduce any potential concerns in the future (although these are not without their drawbacks).

Finally, the failure to detect a cell type in lineage tracing does not necessarily mean that this cell type does not exist. Rather, it may be a reflection of the models used, such as the specific Cre driver or quantification method employed (Rios et al., 2016). Consequently, it may be that MaSCs with a bipotent capacity are involved in another process of mammary gland that was beyond the scope of this work. Nonetheless, these data have highlighted, in multiple

settings, the lineage restriction of the functional stem and progenitor cells that actually contribute to the main morphogenic events in the mammary gland.

4.3.3 Conclusions

We posit that the MaSCs/progenitors that generate the ductal network during puberty are distinct unipotent cells that are distributed throughout the gland during pubertal morphogenesis. Initially found within the TEBs, they proliferate and move towards the subtending duct as it elongates. These unipotent MaSCs/progenitors may then be deposited throughout the ductal network for later recruitment in response to pregnancy hormones to generate alveoli in the lactating gland. Furthermore, embryonic MaSCs also display lineage restriction, as early as E17.5, during development of the rudimentary ductal tree.

Chapter 5

Proliferation in the mammary gland

5.1 Introduction

As discussed in previous chapters, the mammary gland is rare among its epithelial counterparts in that the majority of its development occurs postnatally, with only a minor rudiment developing during embryogenesis. After birth, the rudiment expands isometrically from the initial ducts. The major expansion of the mammary gland then occurs with the onset of puberty, with the surge in ovarian hormones signalling for ductal elongation and branching. After this transformation during puberty, the gland also expands with the onset of pregnancy, with the formation of lobuloalveolar units and side branching during gestation and subsequent functional differentiation during early lactation. With post-lactational regression, known as involution, the gland then returns to a near-nulliparous state. This amazing ability for expansion and contraction can happen multiple times over a lifetime, with bursts of expansion and proliferation followed by subsequent regression and death. Cyclical changes in ovarian hormones during oestrous cycling also results in cyclical side branching and subsequent regression as the gland prepares for pregnancy with each oestrous cycle (Hodson et al., 2013).

Proliferation in the mammary gland has been previously studied by others to try and understand this incredible capacity. Indeed, the TEBs during pubertal development have long be known to be highly proliferative, with tissue sections showing them to be a structure with a high percentage of proliferative cells, particularly in the cap cell region (Paine and Lewis, 2017). Consequently, TEBs have been the proposed as the location of stem cells in the developing gland. However, after pubertal development the gland does not return to quiescence and instead remains dynamic with oestrous cycling. During this period, cyclical increases in progesterone and oestrogen result in secondary and tertiary branching and a process akin to alveolar budding occurs with the gland preparing for pregnancy (Zeps et al., 1999). These then regresses with the drop in progesterone levels at the latter end of the oestrous cycle if a pregnancy this does not occur (Zeps et al., 1999). In the mouse, this cycle takes approximately four to five days, highlighting the ability of the gland to turnover in a short period of time. Consequently, this shows an additional requirement of the presumptive MaSC in gland homeostasis, outside of pubertal and alveolar morphogenesis. Furthermore, identifying regions where proliferation occurs may provide utility in identifying plausible locations / niches for the MaSC in the pubertal and post-pubertal gland.

However, traditionally examination of proliferation in the gland has been carried out in 2D using tissue sections. This both means that investigation of the entire mammary gland is time consuming, but also identifying patterns of proliferation is challenging. Hence, the 3D imaging methods, discussed in chapter 3, provide an approach that allows for a more accurate visualisation of proliferation throughout the entire ductal network, not previously possible through tissue sectioning. Another important consideration is the heterogeneous nature of the gland during different oestrous stages, due to the changes in both morphology and proliferation previously described. Consequently, in order to properly examine the changes in proliferation in the gland at different developmental stages all mice must be normalised for oestrous stage.

5.1.1 Importance

In the previous chapters it was demonstrated that multiple unipotent functional stem cells are responsible for building the ducts in puberty and alveoli in gestation and lactation. However, lineage tracing only demonstrates the fate of these clones. As such, another important characteristic to examine is to ascertain proliferation levels in the mammary gland to appreciate the requirements of these mammary stem and progenitor cells. Similarly, with sustenance of proliferation being one of the original hallmarks of cancer (Hanahan et al., 2011), understanding the proliferation capacity of the mammary gland has important implications for cancer transformation.

5.1.2 Aims

The work described in this chapter aims to define the proliferation levels of mammary epithelial cells during different developmental stages. There are a number of methods that can be utilised to do so, including DNA intercalators that are analogues of DNA and as such can become incorporated into DNA during the S-phase of the cell cycle. After a defined chase period, cells that were proliferating at the time of the pulse (and that have not divided multiple times the since the label) can subsequently be detected (Kee et al., 2002; Zeng et al., 2010). Traditionally, BrdU has been used. However, in recent years its use has been gradually declining as EdU is utilised more frequently. Both molecules are thymidine analogues, however BrdU requires antibody-based detection, which requires use of various harsh chemicals so that the antibody is able to bind to the BrdU which is tightly held within the DNA double helix. EdU on the other hand, can be detected fluorescently using a simple copper-based

chemical "Click" reaction (Zeng et al., 2010). As such, EdU presents as a more useful tool that allows for more co-stains as a consequence of not needing the harsh chemicals.

Accordingly, here both Ki67 staining and EdU incorporation-based examination of proliferation will be used and combined with 3D imaging using optical tissue clearing wholemount confocal microscopy - allowing for an accurate appreciation of the proliferative regions in the gland. Visualisation of proliferating cells will be examined in pre-pubertal, pubertal, young adult and older adult animals

5.2 Results

5.2.1 Oestrous staging optimisation

As discussed above, proliferation in the gland varies with cyclical ovarian hormone changes during the oestrous cycle. Therefore, in order to remove any stage-specific bias mice were oestrous stage-matched. There are several ways to assess oestrous stage in mice that have been described, including monitoring serum hormone levels. However, this is particularly invasive and not feasible to carry out on a routine daily basis. Another technique is to asses cervical cell phenotypes. A number of methods have been developed to do this, such as swabbing of the vaginal entrance, however this can be invasive and risk either inducing pseudopregnancy (Adler et al., 1970; Yang et al., 2009) (which may affect mammary gland biology) or inflammation of the cervical canal (which would confound cytological assessment).

Therefore, we sought to utilise a method that was minimally invasive which would both allow oestrous staging to be carried out on a daily basis and would not have the risk of inducing pseudopregnancy or inflammatory events. This was carried out using small amounts (approximately 50μL) of PBS and a plastic bulb combined with a pipette tip (McLean et al., 2012). The PBS can then be flushed in and out of the vagina via aspiration, collecting exfoliated cells, meaning no insertion into the vaginal canal is needed. Lavage samples are then placed onto histology slides, dried and subsequently stained using a simple toluidine blue dye. Slides can be immediately assessed without the need for fixation or mounting and so consequently this process is very time efficient.

Oestrous stage is assessed based on the relative ratios of three different cell types in the vaginal cytology samples: nucleated epithelial cells, cornified squamous epithelial cells, and leukocytes (Figure 5.2.1a). This reflects changes in cervical topology that occurs with the different endocrine environments during the different phases of the oestrous cycle. During proestrous, cells are mostly clusters of round, well-formed nucleated epithelial cells (Figure 5.2.1b,i). Next, during oestrus, cells are predominantly cornified squamous epithelial cells (Figure 5.2.1b,ii). During metoestrous, small leukocytes are detected mixed with both cornified squamous and nucleated epithelial cells (Figure 5.2.1b,iii). Finally, during dioestrous rare cornified squamous epithelial cells may still be present, however mostly leukocytes are detected (Figure 5.2.1b,iv).

Consequently, all mice were oestrous staged at the same time each day, 7 days prior to being assessed for proliferation. The reason for this is two-fold: firstly, all experiments were carried out when the mice were in dioestrous, when proliferation is at its highest due to increased progesterone levels (Giraddi et al., 2015; Shehata et al., 2018). Secondly, this allowed for identification of mice which had incorrect cycling – those that did not correctly cycle were therefore not used for this study. A normal oestrous cycle was defined as following these defined stages over 4 – 5 days.

To assess the proliferation of the mammary gland at different developmental stages four different time points were examined: 4 weeks for pre/peri puberty, 7 weeks for late puberty, 12 weeks for young adult and 24 weeks for older adult. CUBIC optical clearing (the supplementary R1a protocol) was utilised in order to gain better visualisation due to superior depths that can be achieved with the protocol, which is particularly relevant when analysing older, larger glands.

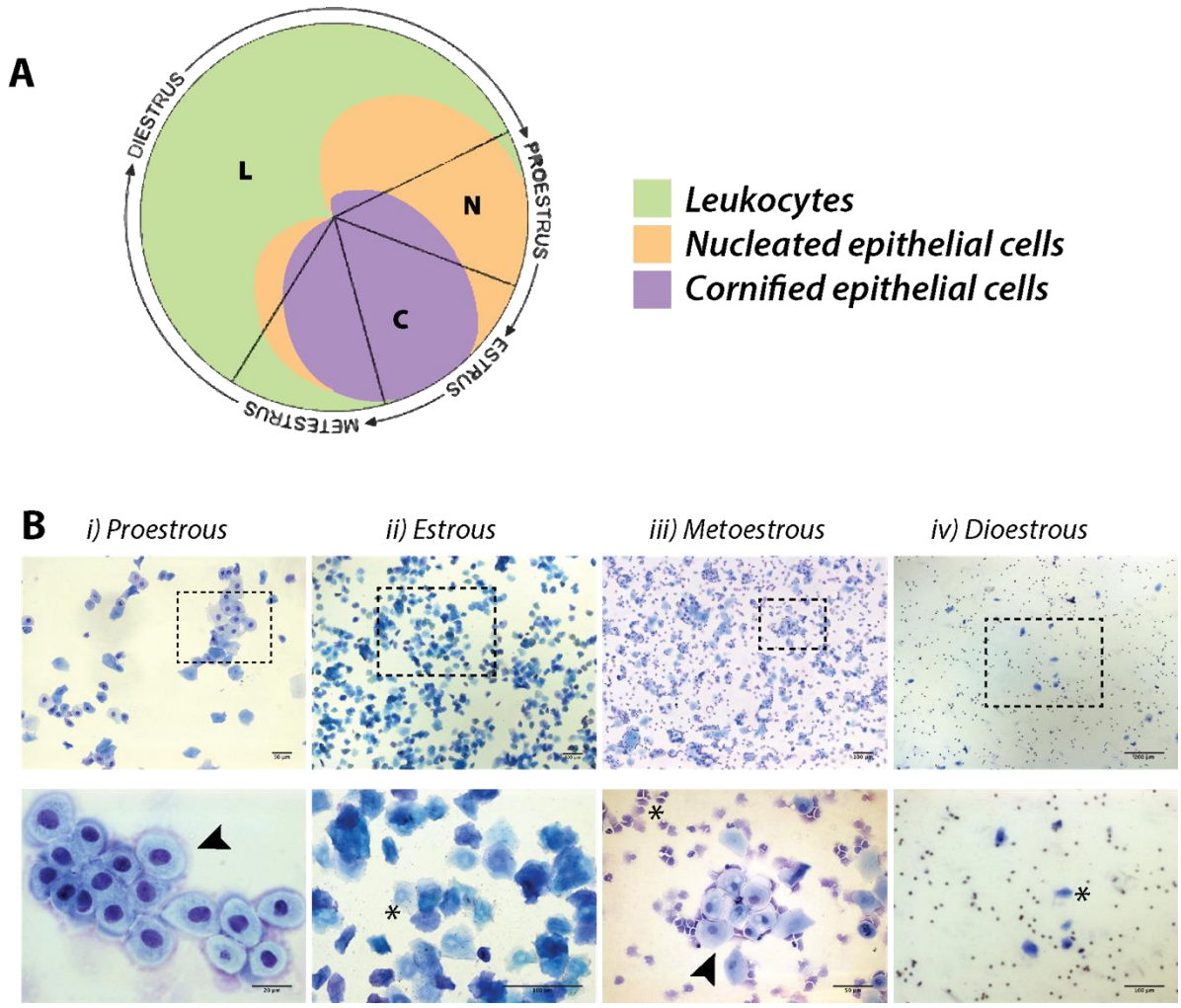


Figure 5.2.1 Outline of method for determining oestrous stage in mice using vaginal lavage

(A) Schematic outlining the differing cellular compositions of vaginal lavages during the four different oestrous stages, based on the relative amounts of leukocytes (L), nucleated epithelial cells (N) and cornified epithelial cells **(C)**. Diagram based on (Cora et al., 2015). **(B)** Light microscopy images showing representative vaginal cytologies obtained during i) proestrous, ii) oestrous, iii) metoestrous and iv) dioestrous. Nucleated epithelial cells highlighted by arrow heads and cornified epithelial cells highlighted by asterisks.

5.2.2 Ki67

Proliferation can be visualised in a number of ways, with Ki67 staining widely used as one of the easiest ways to detect cycling cells that are in any stage of the cell cycle, excluding G0 (Scholzen and Gerdes, 2000). Ki67 staining in the mammary gland has been previously investigated using 2D tissue sections (Zeps et al., 1999; Fernandez-Gonzalez et al., 2009), and therefore it was decided to use Ki67 as an initial investigation of proliferative regions that can be examined in the mammary gland of optical cleared tissue. Accordingly, mammary glands were staining for Ki67 and cleared using CUBIC-based optical clearing. Small intestine samples were taken as a proliferative control (Figure 5.2.2a), in addition to no-primary control staining to identify non-specific or background staining (Figure 5.2.2b)

At 4 weeks, the gland is composed of a basic ductal structure with approximately 5 main branches. The general consensus is that puberty begins in a mouse at approximately 3 – 4 weeks of age; however like humans, this number can vary from mouse to mouse based on several factors. Consequently, this age encompasses mice in either pre-puberty or early puberty, depending on whether they have been exposed to the pubertal hormones that instruct that gland to expand (Hinck and Silberstein, 2005).

As expected, proliferation was seen concentrated in the cells of the TEBs in 4 week mammary glands. Largely this was seen in the outer layers of body cells and within the SMA+ cap cells (Figure 5.2.2a). However, in some cases SMA+ cap cells could be detected within the body of the TEB, some of which co-stained with Ki67 (Figure 5.2.2b, arrow head). Previously, it was hypothesised that cap cells represent a population of bipotent stem cells that are able to produce cells of both epithelial lineages, and they do so by entering the TEB body and laying down the cells in the subtending duct. Conversely, excess cap cells have been said to enter the TEB body to undergo programmed cell death (Paine et al., 2016) (see chapter 1 for discussion of these hypotheses).

Furthermore, Ki67 staining was also observed in the subtending ducts leading up to the TEBs, both at branch points (Figure 5.2.4a) and further away from the TEB (Figure 5.2.4b, arrow head). Traditionally, proliferation during puberty is thought to be led by the TEBs. These areas in the duct have been reported to remain quiescence once laid down by the expanding TEB, until required to form lobuloalveolar expansion during pregnancy, and to a lesser degree with oestrous cycling.

Similar observations were also seen in the mammary glands of 7 week mice. The majority of Ki67 labelled cells were seen in the TEBs of the mammary gland (Figure 5.2.5a,b) but again Ki67 staining was detected also further down the ductal network (Figure 5.2.6a,b). Moreover, stromal cells around the ducts also expressed Ki67 (Figure 5.2.6b, arrow heads). Definitive identification of these cells would require further investigation but given the role of stromal cells in pubertal gland development (discussed in introduction) this result is not surprising.

Within post-pubertal mammary glands, Ki67 staining was much more infrequent. Indeed, uncommon rare cells could be detected, but the majority of staining was seen in the stromal compartment both at 12 weeks (Figure 5.2.7 and Figure 5.2.8) and at 24 weeks (Figure 5.2.9 and Figure 5.2.10). This reduction in proliferation in the post-pubertal mammary gland is expected given the fact that ductal morphogenesis is complete. However, given proliferation is still required for expansion in oestrous, small amounts of proliferation are still expected.

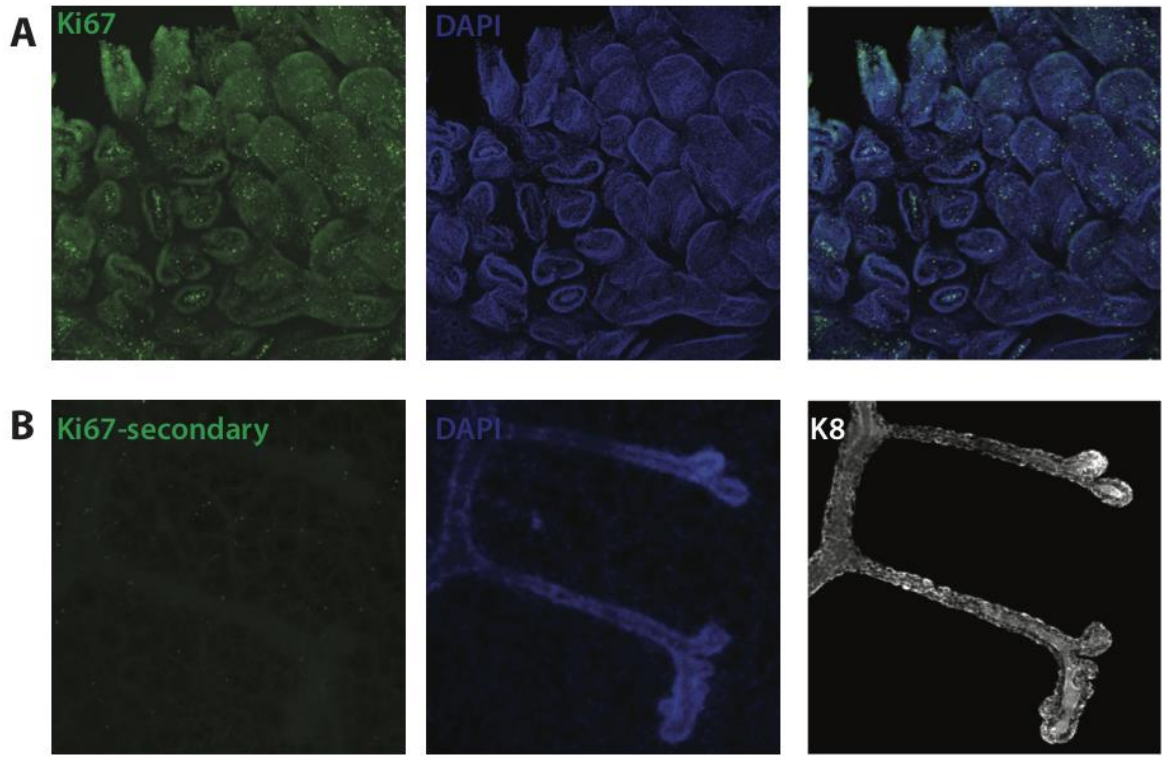


Figure 5.2.2 Ki67 staining controls in the murine intestine and mammary gland

(A) Sections of the small intestine were used as proliferation controls to ensure staining was compatible with optical clearing **(B)** Representative images of no primary control staining, with tissue only labelled with K8 primary antibody, followed by K8 and Ki67 secondary antibodies and DAPI used for the nuclear stain. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

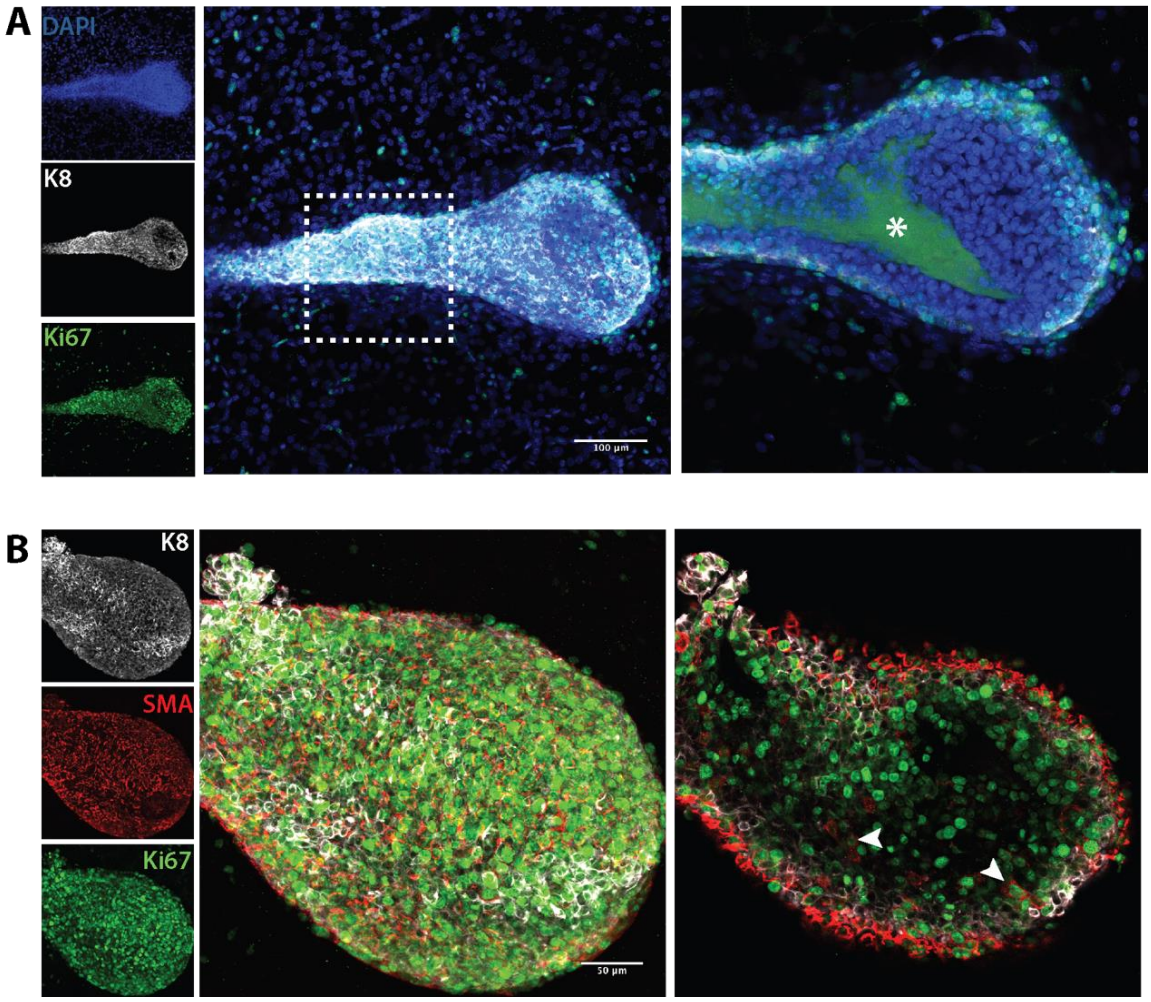


Figure 5.2.3 Ki67 staining in 4 week mammary glands

4 week mammary gland tissues were stained with K8 and Ki67, with DAPI used for nuclear staining. Ki67 staining was seen in both larger and smaller TEBs demonstrating that they were highly proliferative. **(A)** and **(B)** show representative images showing SMA+ cap cells in the body of the TEB co-staining with Ki67. **(B)** Asterisk marks CUBIC-related autofluorescence (see chapter 3 for details). Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

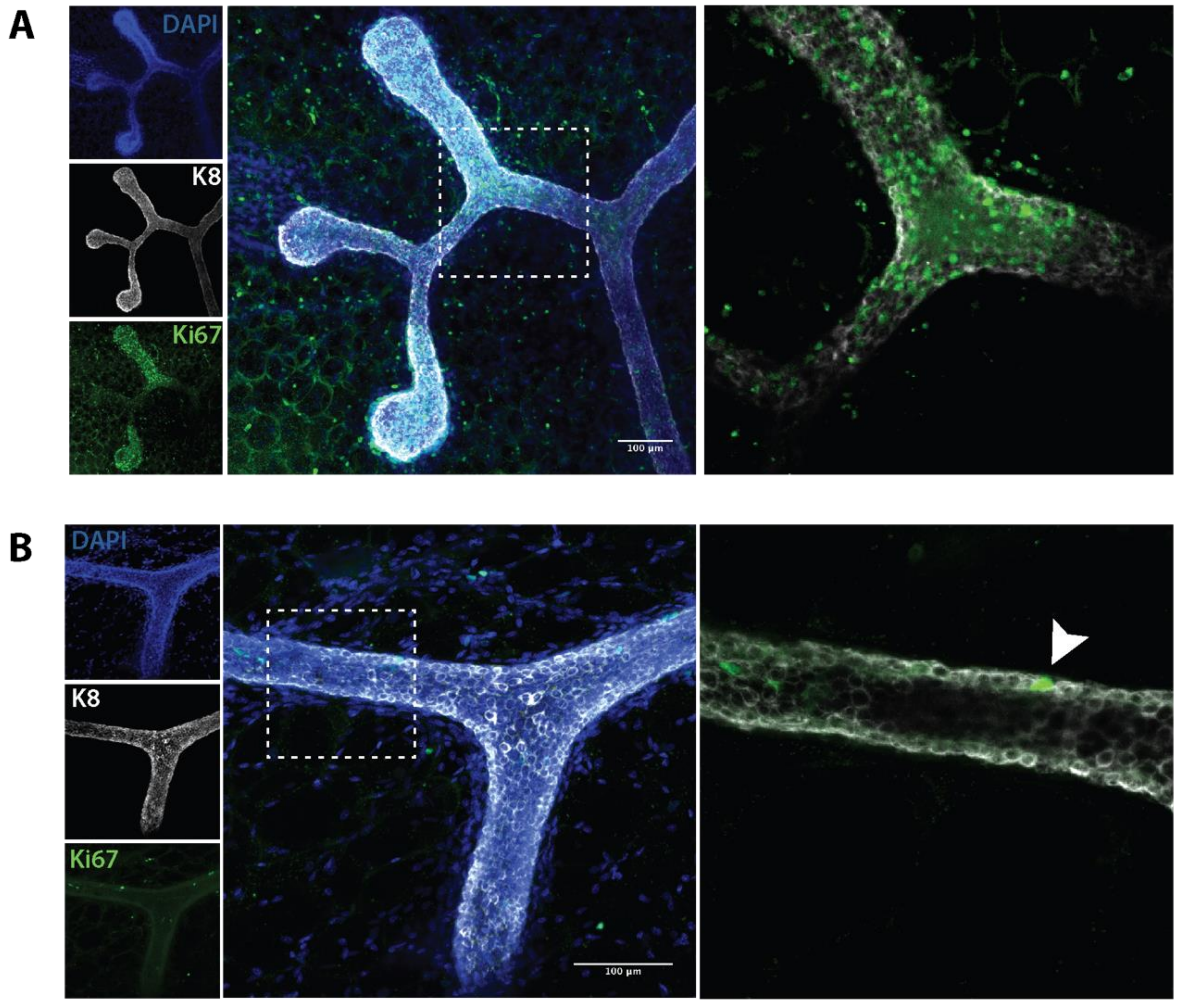


Figure 5.2.4 Ki67 staining seen in ductal regions of 4 week old mammary glands

(A) Proliferation seen by Ki67 staining in multiple regions within a trifurcated TEB, including further down the subtending duct. Note the on/off nature of Ki67 staining, with the middle TEB staining less than the others. **(B)** Representative image showing more sparse proliferation in cells further down the ductal region. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

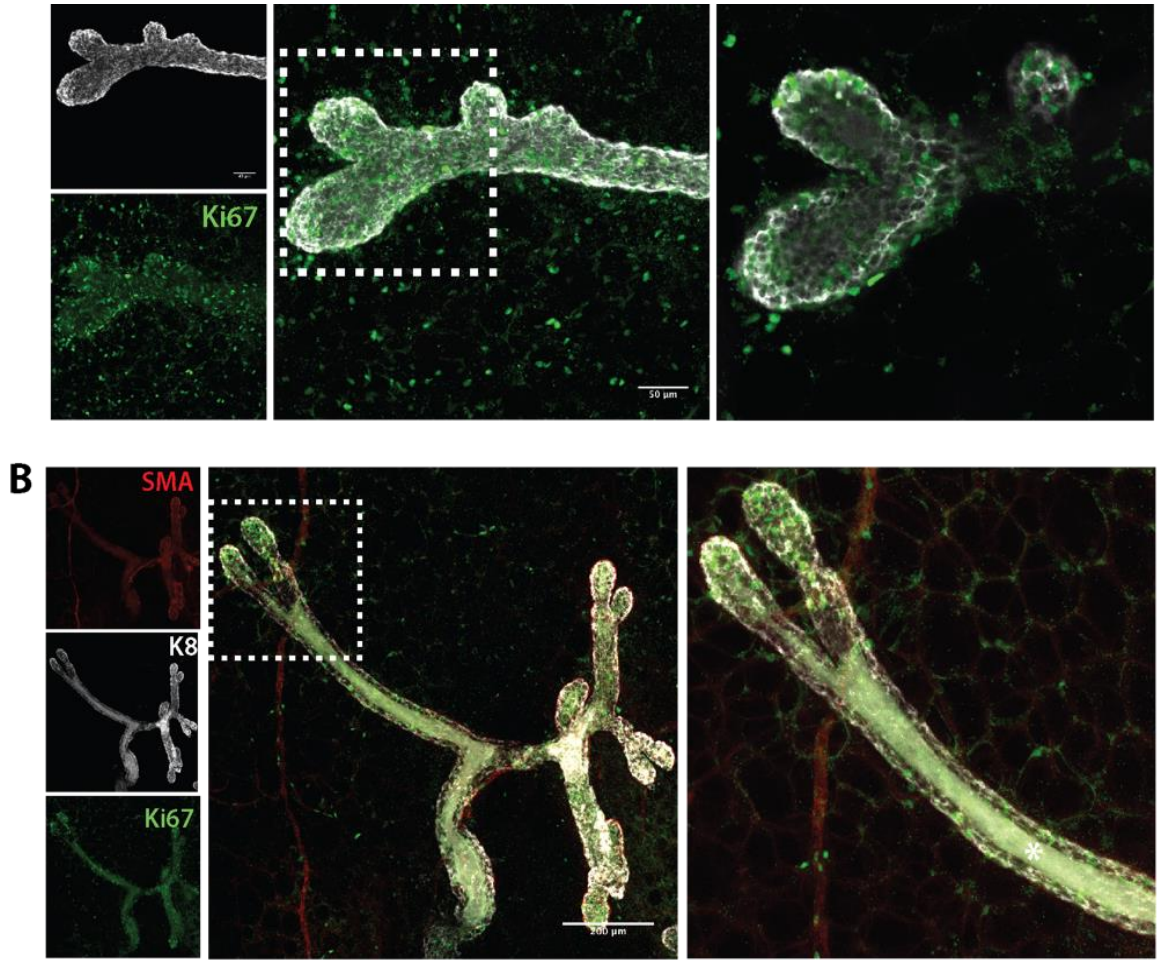


Figure 5.2.5 Ki67 staining in terminal end ducts of 7 week mammary glands

(A) and **(B)** show representative images of proliferation shown by Ki67 staining in bifurcating TEBs of 7 week mammary gland. **(B)** Asterisk marks CUBIC-related autofluorescence (see chapter 3 for details). Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

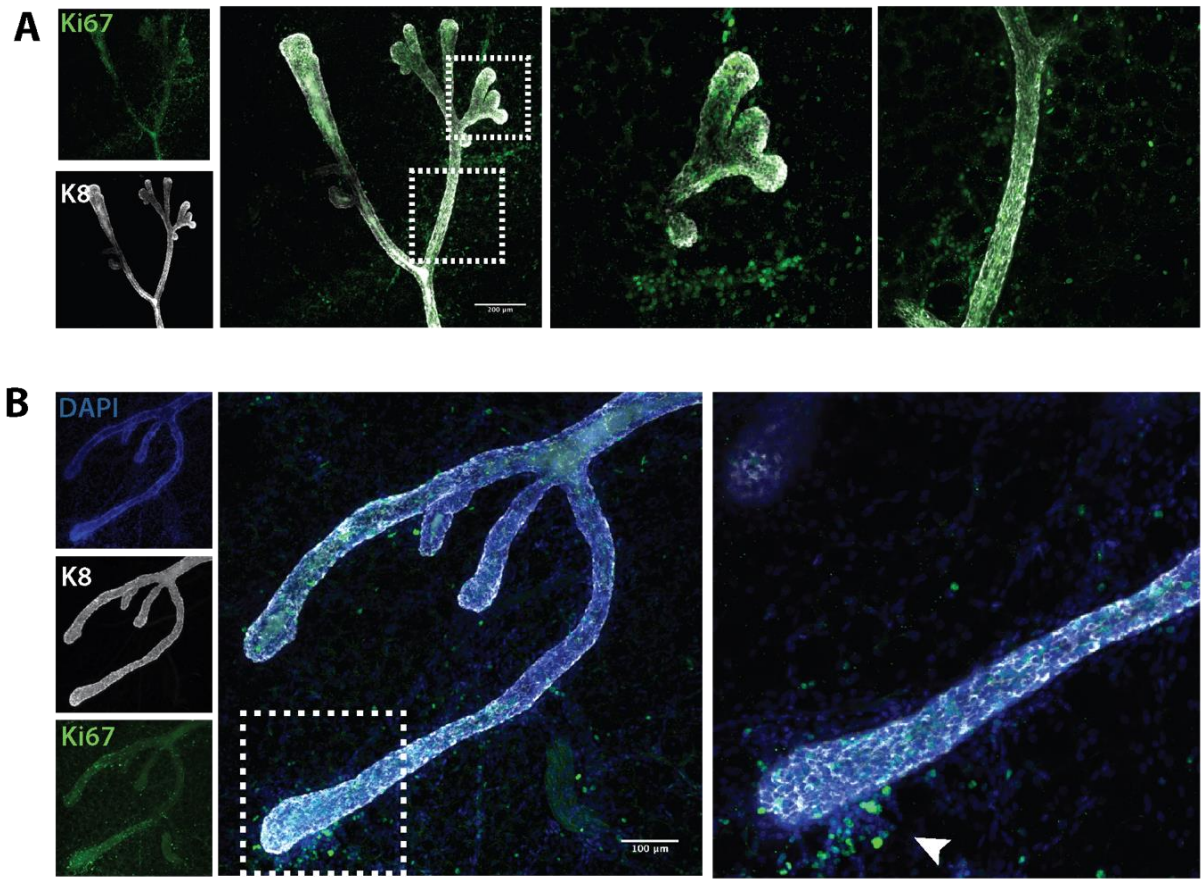


Figure 5.2.6 Ki67 staining in 7 week mammary glands

Representative images showing that not all areas are proliferative in 7 week mammary glands. **(A)** shows a ductal region with proliferation in two areas. **(B)** representative images showing Ki67 staining in stromal cells surrounding terminal end ducts (arrow head). Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

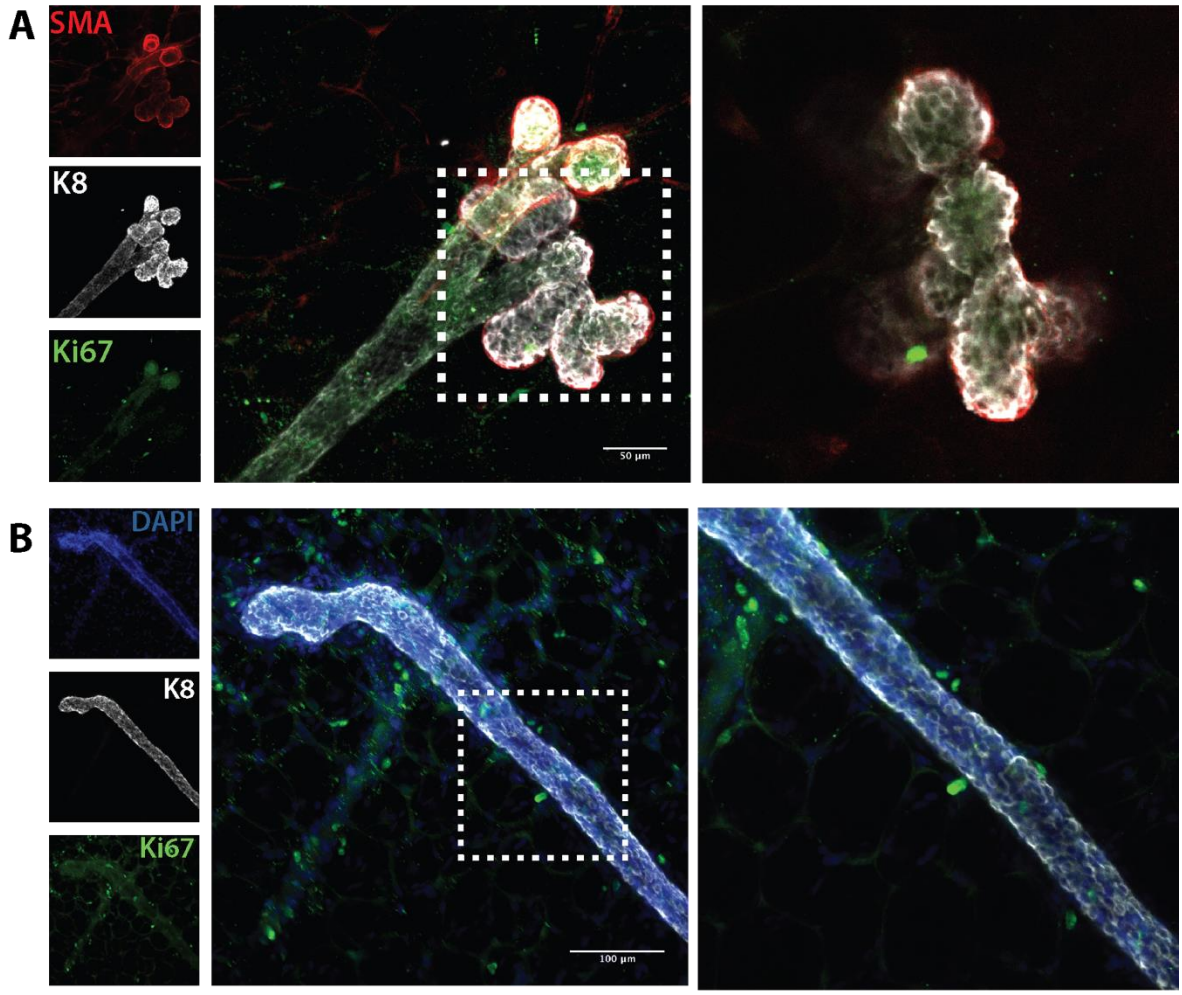


Figure 5.2.7 Ki67 staining in 12 week mammary glands

Representative images showing rare proliferative cells in 12 week mammary glands. **(A)** Example of budding region at a terminal end bud **(B)** Ki67+ proliferating cells in both ductal and stromal cell types. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

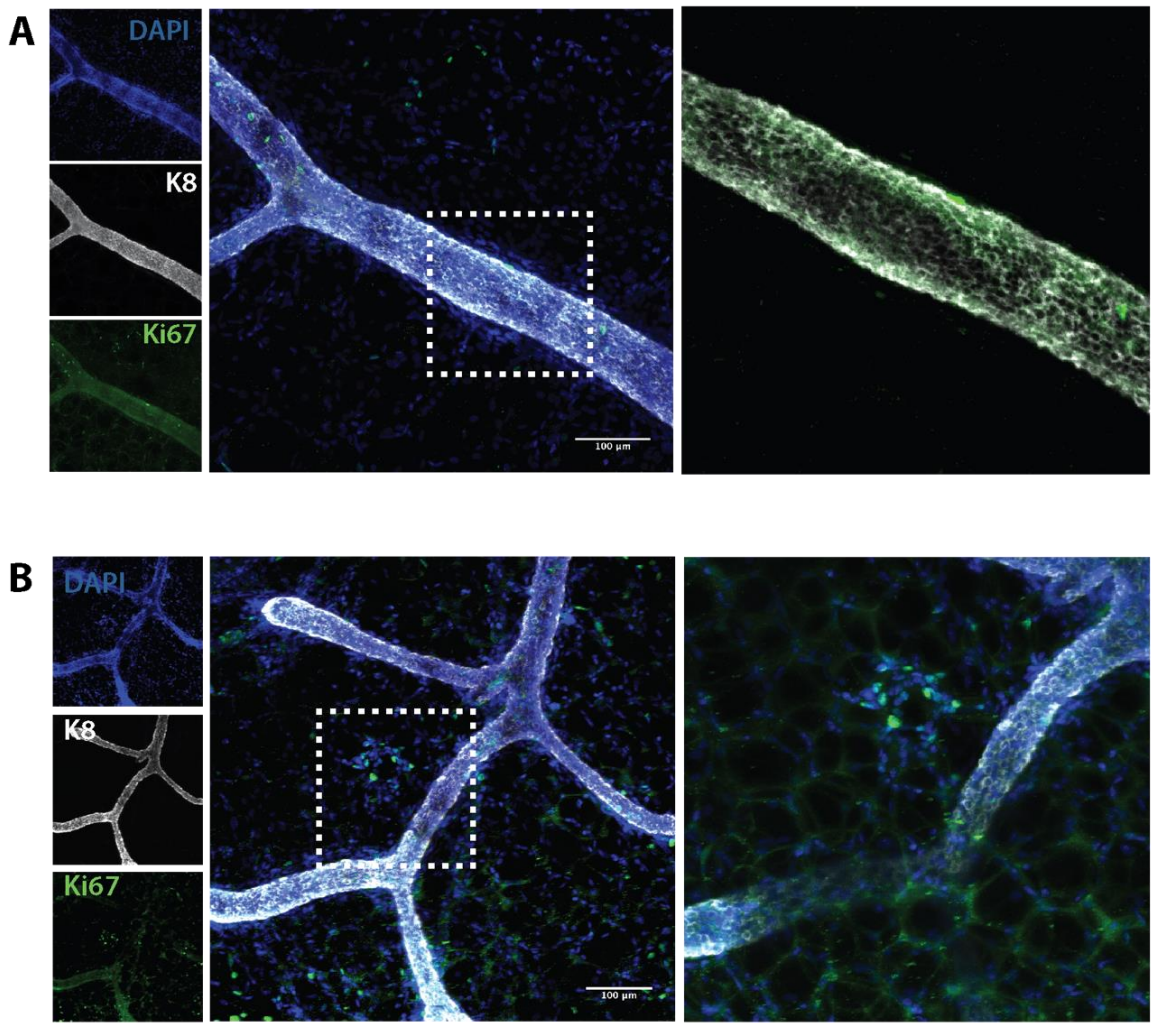


Figure 5.2.8 Ki67 staining examples in 12 week mammary glands

Further examples of Ki67 staining in ductal **(A)** and stromal **(B)** regions of 12 week mammary glands. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

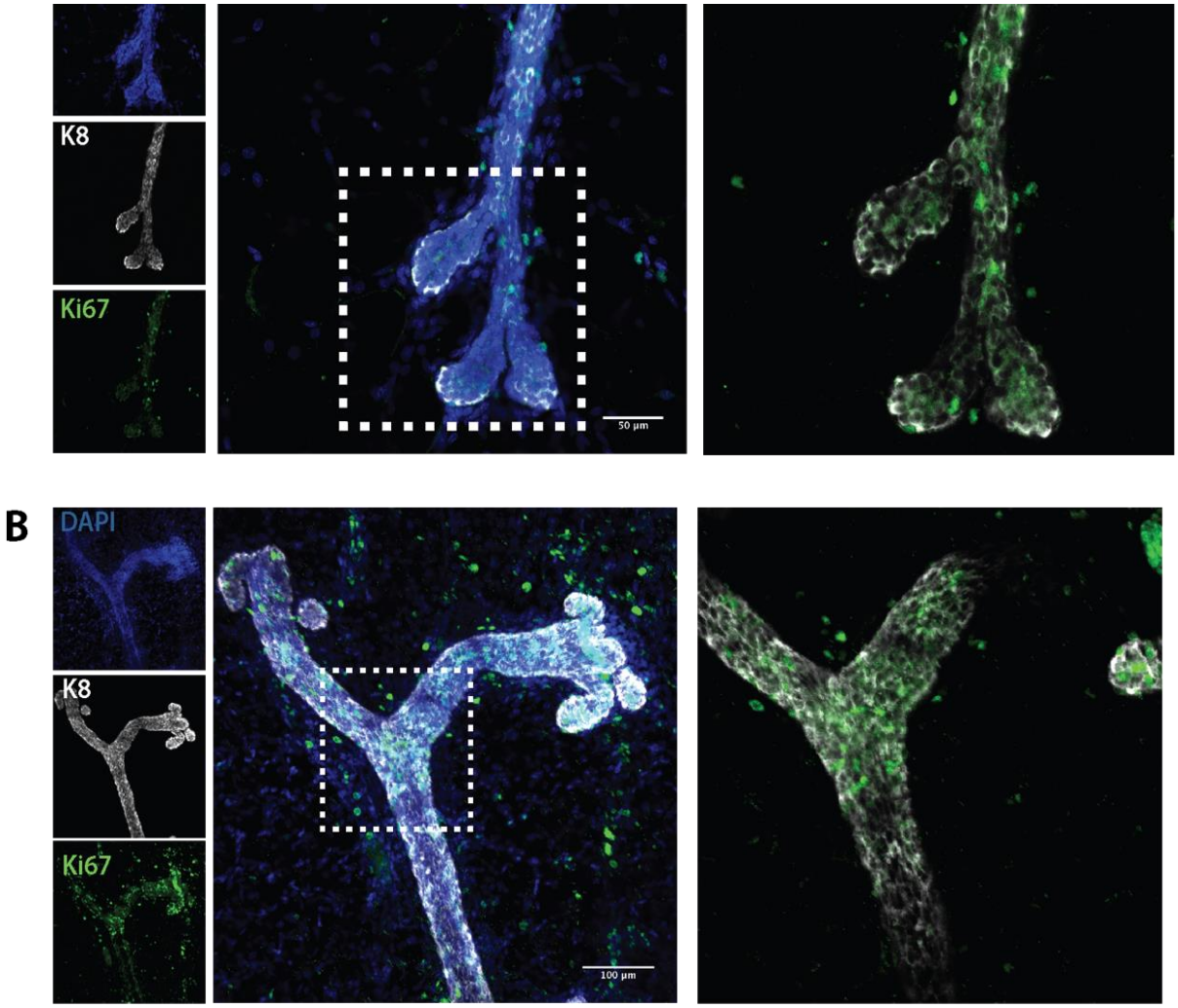


Figure 5.2.9 Ki67 staining in 24 week mammary glands

Representative images showing Ki67 staining in terminal ducts of adult 24 week mammary glands. **(A)** and **(B)** show Ki67+ proliferating cells in both ductal and stromal cell types. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

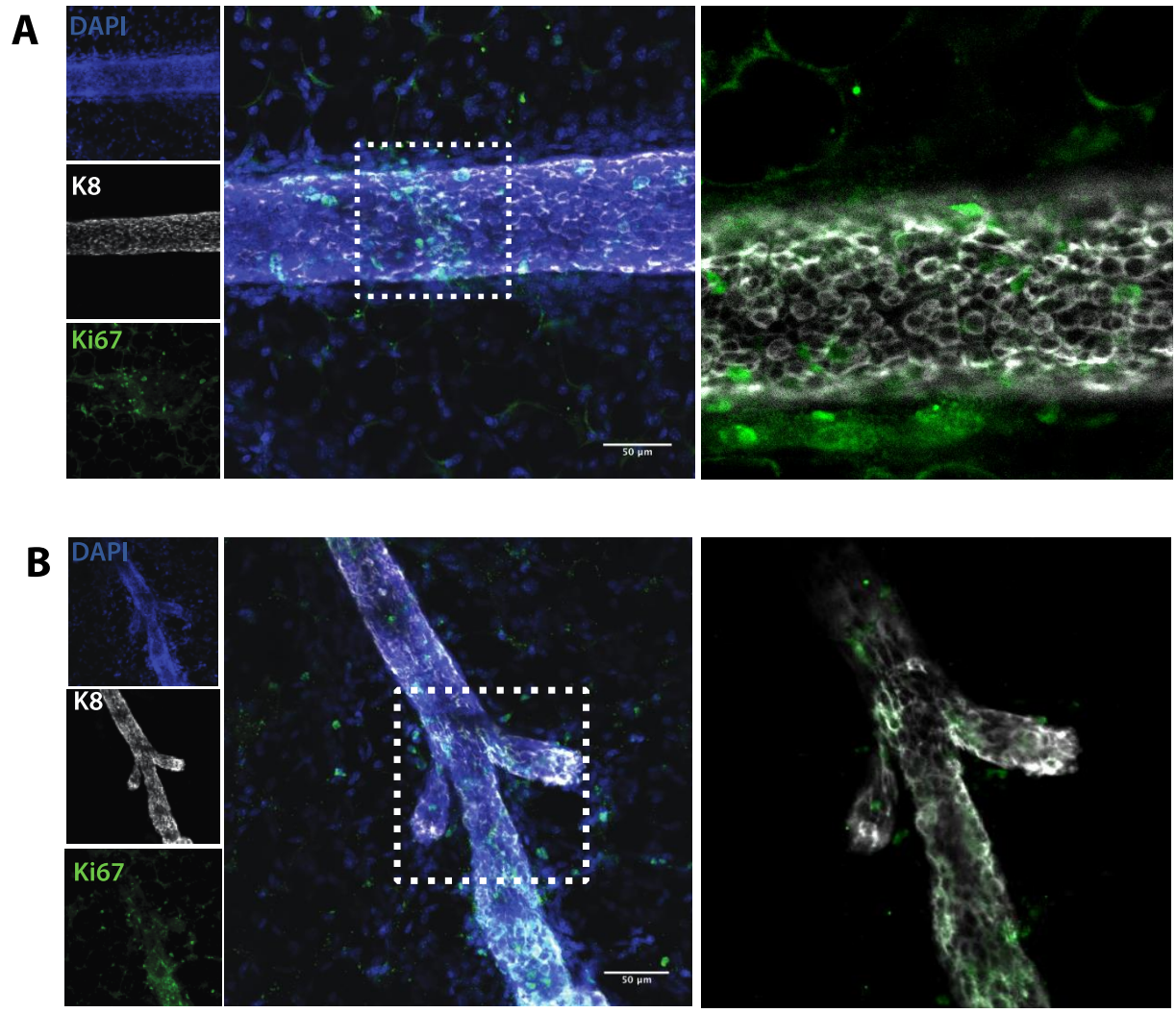


Figure 5.2.10 Ki67 staining in 24 week mammary glands

Representative images showing Ki67 staining in ductal regions of adult 24 week mammary glands. **(A)** and **(B)** show Ki67+ proliferating cells in both ductal and stromal cells. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

5.2.3 EdU

Ki67 staining to examine proliferation allows identification of cells within any stage of the cell cycle. However, use of DNA intercalators allows for a more precise measurement of active cell division in the mammary gland throughout development. Indeed, EdU incorporation only occurs during the S-phase of the cell cycle, as opposed to all the stages that are labelled with Ki67. As such, this allows for detection of cells that have actively proliferated within the chase period. Here, a 4-hour pulse chase period was used as an optimal amount of time to specifically mark actively proliferating cells (for example rather than investigating label retention). Following this, tissues were stained with basal and luminal cell markers and cleared using either SeeDB or CUBIC optical tissue clearing, with non-injected control and intestine samples for injection and staining controls (Figure 5.2.11).

5.1.1.1 Pre-pubertal EdU proliferation

Similar to tissues examined with Ki67 staining, expanding ductal trees with TEBs could be visualised in some mice. As expected based on both previous work in the field (Kenney et al., 2001; Fernandez-Gonzalez et al., 2009; Paine et al., 2016) and data from Ki67 staining, high levels of EdU incorporation was seen in TEBs. This was observed in both large bifurcating TEBs, and smaller structures (Figure 5.2.12, a and b respectively). When analysing the composition of proliferation in the TEBs, both cap and body cells could be seen to have incorporated EdU.

Proliferative SMA+ cells, which were presumably cap cells, could again be detected within the body of the TEB in a number of cases, as was seen with Ki67 staining. (Figure 5.2.12c, arrowheads). These cells have been described to be undergoing programmed cell death, but previous reports have also postulated a role for these cells in formation of the ductal lumen (Kenney et al., 2001). Further staining with a marker of cell death marker, such as CC3, would provide more insights into the role of these cells within the TEB. Interestingly, EdU incorporation seemed to be binary - in that TEBs were either "on" or "off", with either most cells proliferating or none. This phenomenon also occurred within bifurcated TEBs (Figure 5.2.13). Indeed, this has been seen using other methods of wholemount imaging (Scheele et al., 2017), and has been suggested to be a mechanism to ensure the fat pad is evenly filled and to prevent ductal clashing.

Furthermore, EdU proliferation could also observed within the main duct proximal to the nipple opening of some mammary glands (Figure 5.2.14). This is interesting considering this area

develops and expands during embryonic development and to a lesser degree after birth. However, given the significantly larger diameter of this duct in comparison with others, it may be that these proliferating cells are responsible for expanding the duct latitudinally, providing a strong and reinforced duct around the nipple for milk expulsion. Indeed, this could be seen by the multiple layers of K8+ luminal cells rather than a single layer (Figure 5.2.14a, inset). 3D imaging also revealed proliferation in stromal cells surrounding this area, including a vast network of SMA+ mesenchymal cells surrounding the nipple opening (Figure 5.2.14a). Further staining of this structure would need to be done to further identify these cells, but these may be either smooth muscle cells or blood vessels, both of which would be marked by SMA.

Use of optical tissue clearing allowed wholemount imaging of the mammary glands which meant visualisation of entire regions of the ductal tree could be achieved. This revealed that proliferation was not always confined to the TEBs, in contrast to what has been previously shown. Confirming the observation seen with Ki67 staining, proliferation could also be seen in zones throughout the ducts, in both the basal and luminal cells (Figure 5.2.15a,b). In some of these regions there appeared to be a slight distention of the duct (Figure 5.2.15b, inset) - this may potentially suggest that these are regions where secondary lateral branches will form, but this is not possible to confirm without intravital imaging. However, this still shows for the first time that the ducts can remain proliferative, possibly to facilitate side branching or ductal elongation. Indeed, buds on the side of branches were also seen to be highly proliferative (Figure 5.2.16).

Another interested morphological finding was detected. Within a number of ducts, small K8 labelled cells could be detected. These were significantly smaller than normal luminal cells and appeared within the lumen (Figure 5.2.17b, arrowhead). The morphology and location of these cells is reminiscent of a population of Notch expressing cells that were identified using Notch reporter mice, termed "S" cells. These cells were seen to be arranged in "strings" surrounding a population of larger, Notch-expressing cells, termed "L" cells. Indeed, measurement of the smaller population of cells identified here showed they were on average 3.02 μm ± 0.49 in diameter – similar in size to the previous described population (2.80 μm ± 0.25) (Šale et al., 2013). Further work involving staining for Notch activity, and most importantly functional studies would be required to validate this definitively.

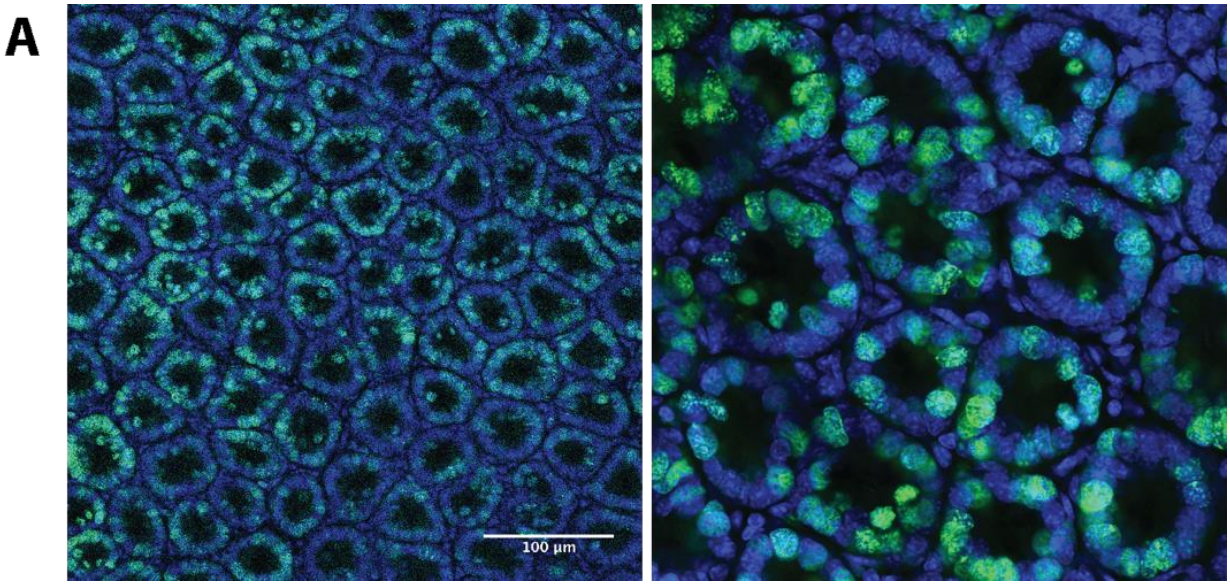


Figure 5.2.11 EdU proliferation seen in small intestine control tissues

25mg/kg EdU was injected into 4, 7, 12 and 24 week-old-mice in dioestrous followed by a 4-hour chase. Representative images show proliferative region of the small intestine, representing both EdU injection and staining reaction control.

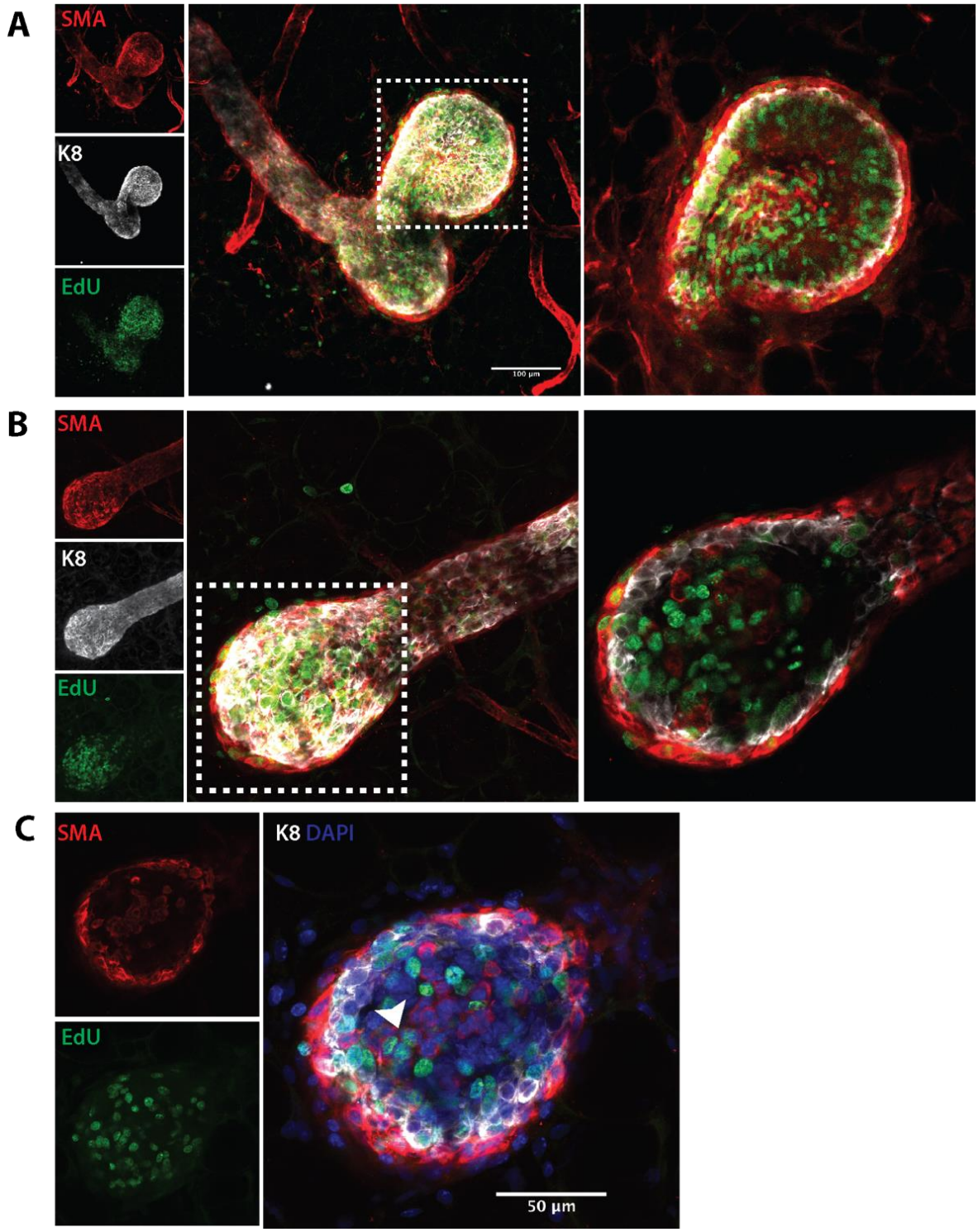


Figure 5.2.12 EdU proliferation in 4 week mammary glands terminal end buds

High levels of proliferation can be seen in the TEBs of early pubertal mammary glands. **(A)** shows representative image of EdU incorporation in a bifurcating TEB, with proliferation seen in both cap and body cells. **(B)** Representative image showing a smaller TEB with EdU incorporation, including in SMA positive cells within the body of the TEB, with magnified z-section of boxed area shown in **(C)**. Arrow head shows SMA+ EdU+ cells.

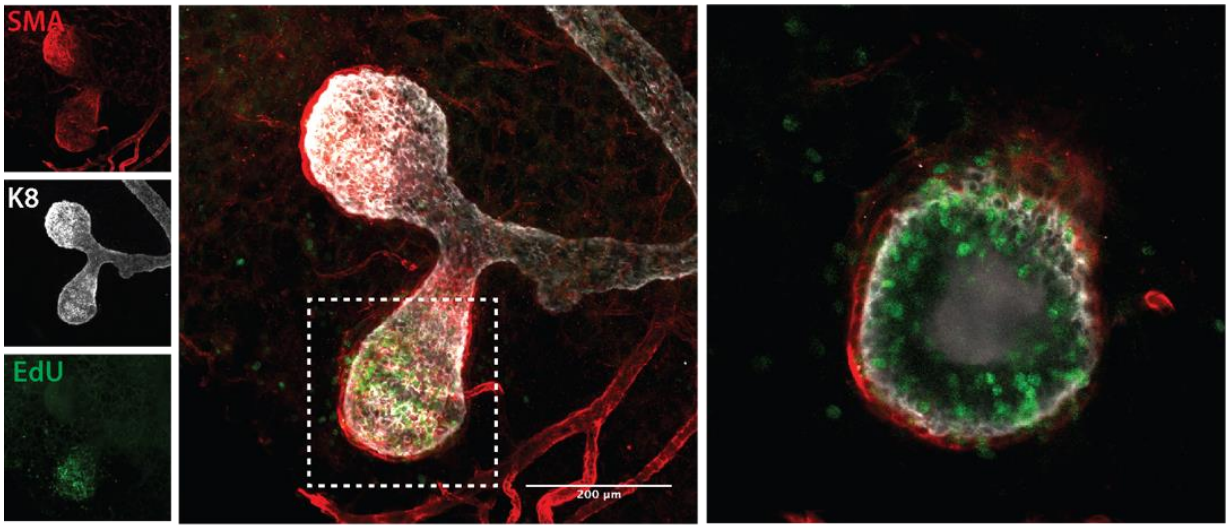


Figure 5.2.13 EdU incorportation in a bifucated TEB of 4 week mammary gland

Representative image of a birfucated TEB with binary proliferation seen in only one half of the TEB. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

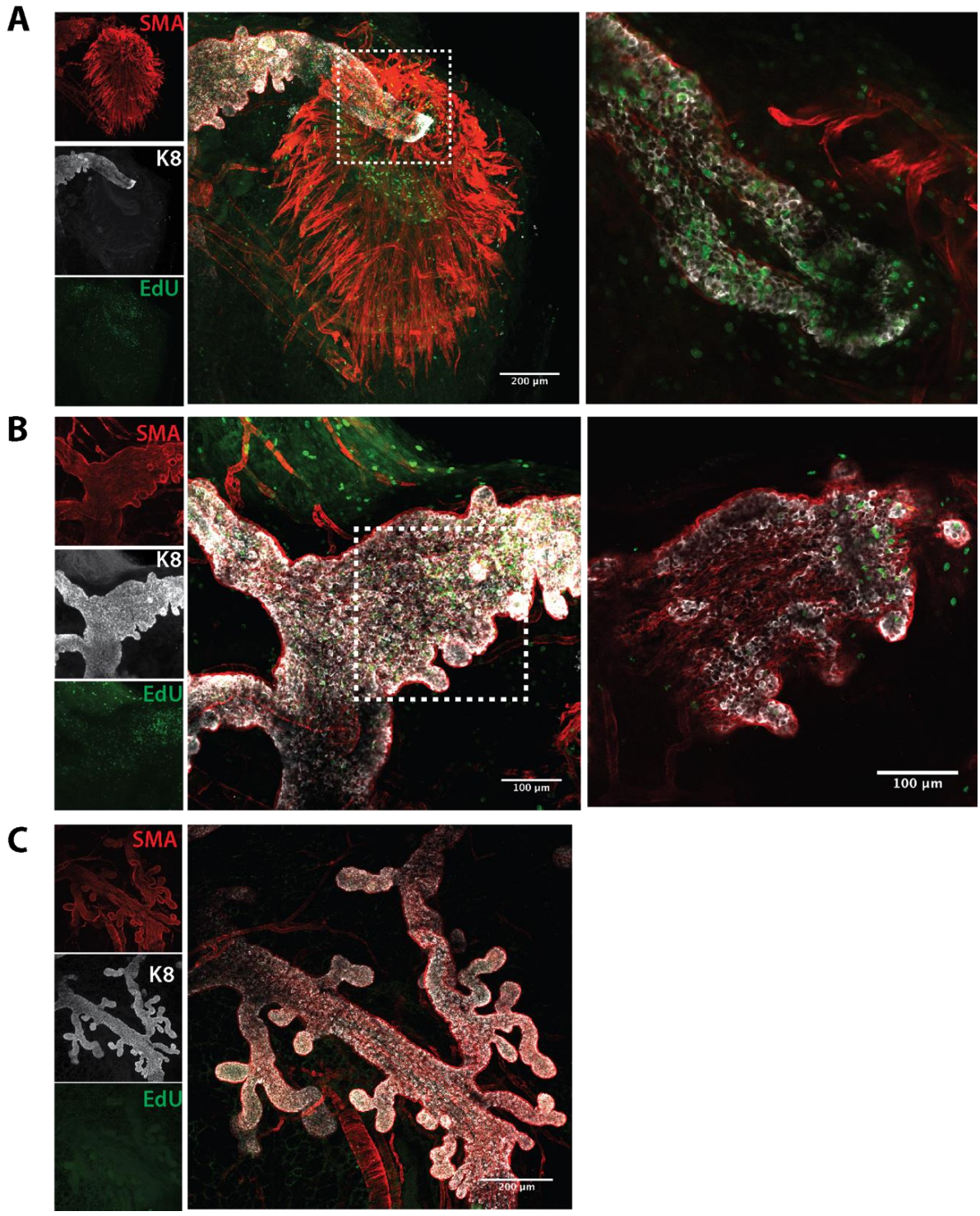


Figure 5.2.14 EdU incorporation proximal to the nipple in 4 week mammary glands

(A) Proliferation is seen in the nipple opening region in the initiating duct in a 4 week mammary gland. SMA+ mesenchyme is also shown with EdU incorporation. **(B)** Proliferation seen in regions further into the ductal network, which is not consistently the case **(C)**.

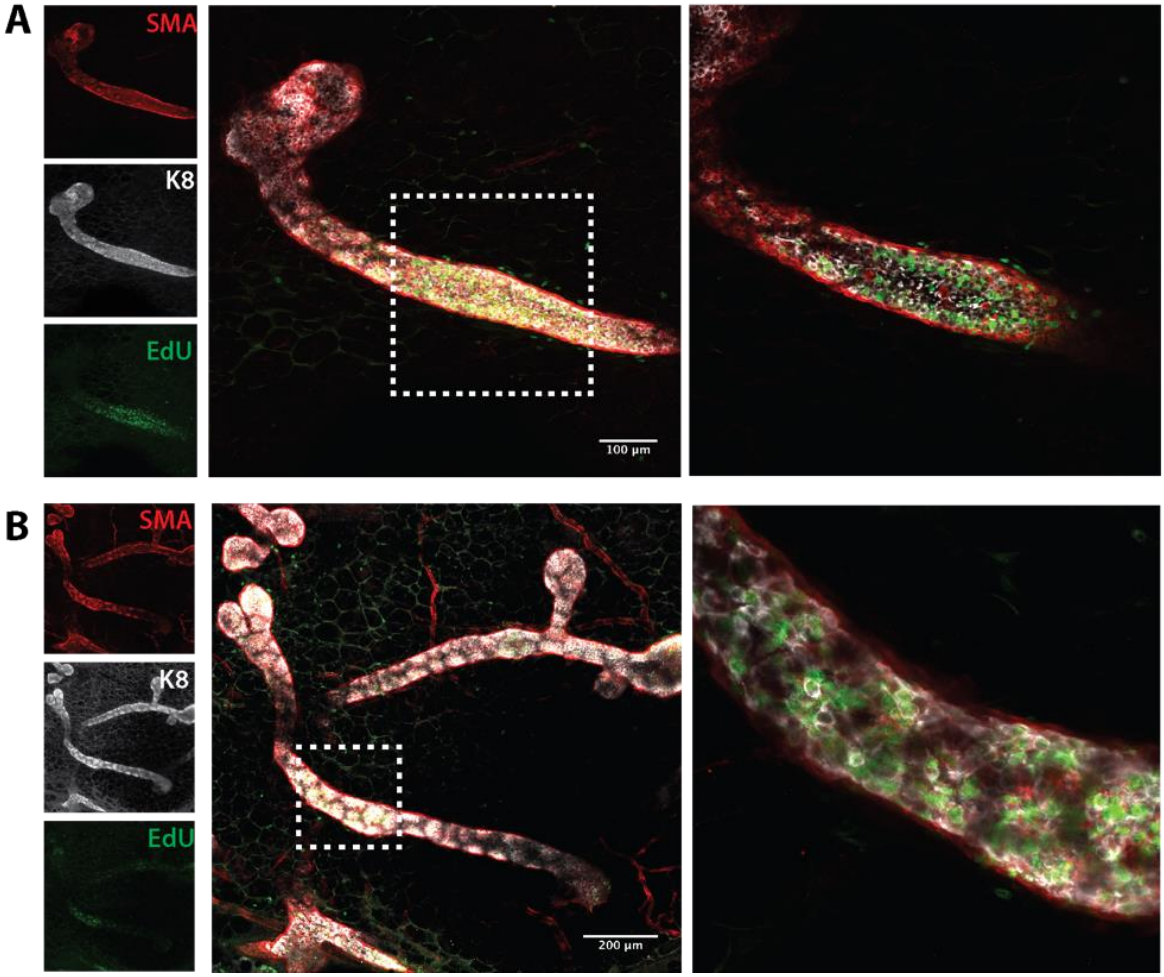


Figure 5.2.15 EdU incorporation seen in the ductal network of 4 week mammary glands.

(A) and **(B)** show representative examples of EdU incorporation seen in ductal regions during early puberty, in both basal and luminal cells. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

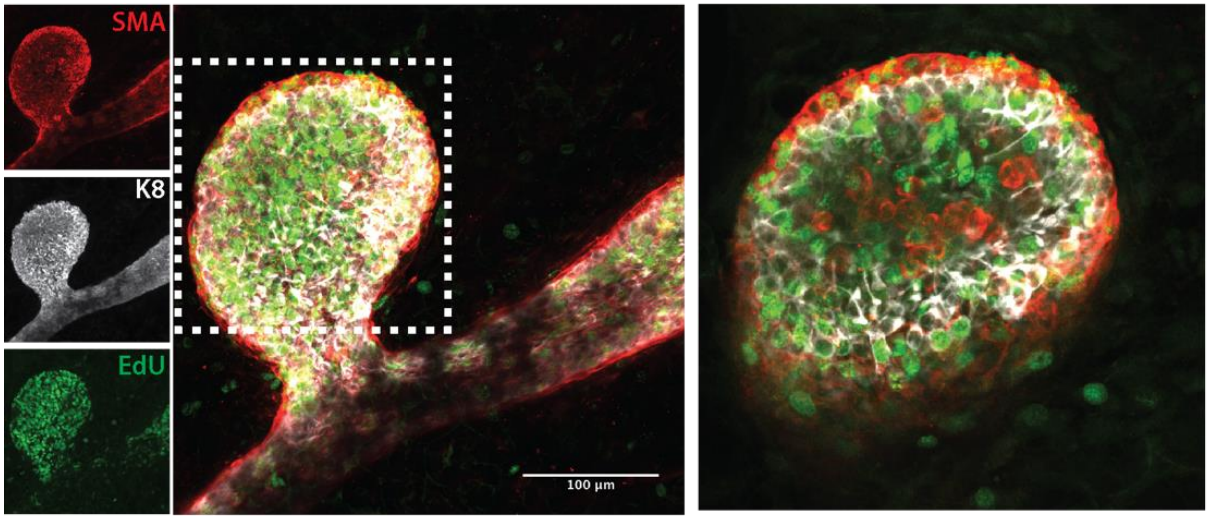


Figure 5.2.16 EdU incorporation seen in the side buds of 4 week mammary glands.

Representative example of EdU incorporation seen in side buds during early puberty, in both basal and luminal cells. Images show maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

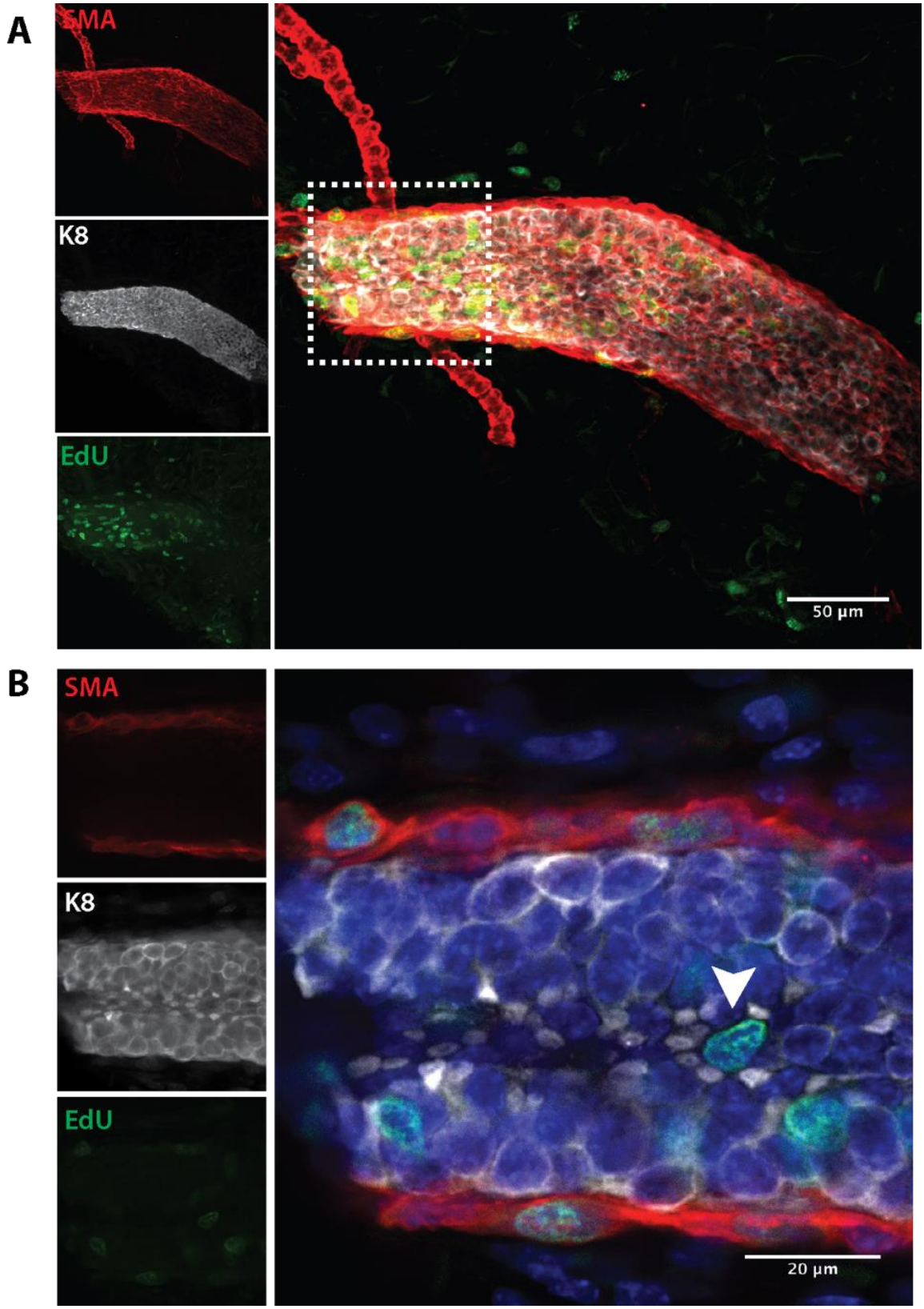


Figure 5.2.17 Population of K8 cells seen in 4 week pubertal mammary glands

(A) Representative images of ductal proliferation seen in both basal and luminal cells of 4 week mammary glands. **(B)** Magnified region of **(A)**, showing a population of small K8+ cells surrounding a larger EdU+ cell.

5.2.1.1 Pubertal proliferation

Mammary glands from 7 week old mice where next examined to investigate proliferation in the latter phase of pubertal development. At this stage, some TEBs remain at the tips of ducts that have not reached the extremities of the fat pad, which is yet to be completely filled. Consequently, proliferation could still be seen in some remaining TEBs (Figure 5.2.18a,b,c). Examination of cell types demonstrated EdU incorporation in both cap and body cells of the TEBs. However, proliferation was much more heterogenous, with overall less proliferation than the younger time point examined. The binary nature seen in 4 week mammary glands was seen again at this age, with not all TEBs and terminal ducts showing high levels of proliferation – indeed, a number of terminal end ducts and side branches has little to no EdU incorporation (Figure 5.2.19a,b)

Interestingly, at this age high levels of proliferation were also seen in the stromal cells surrounding TEBs (Figure 5.2.20a,b,c). The importance of stromal cells in providing both signalling cues and structural support has been previously reported, as discussed in the introduction (Williams and Daniel, 1983; Gouon-Evans et al., 2000; Lilla and Werb, 2010). Further staining would be required to definitively identify which stromal cells these may be, but cells including macrophages, mast cells and fibroblasts have been implicated in supporting ductal growth. It is not fully understood if these cells proliferate *in situ* or home from elsewhere, so this finding suggests that cues within the mammary stroma may induce expansion of these stromal cells.

Similar to early puberty, ductal proliferation was also seen in some areas of 7 week mammary glands (Figure 5.2.21a)., although not to the same extent as that seen in earlier pubertal stages. Indeed, at this age proliferation appeared to be much more heterogenous, with some areas containing no EdU labelling – even within TEBs (Figure 5.2.21b), and expectedly in mid-ductal regions (Figure 5.2.22). Presumably this may be due to the fact that the fat pad is much more filled at this age and as such not as much proliferation is required to expand the ductal network.

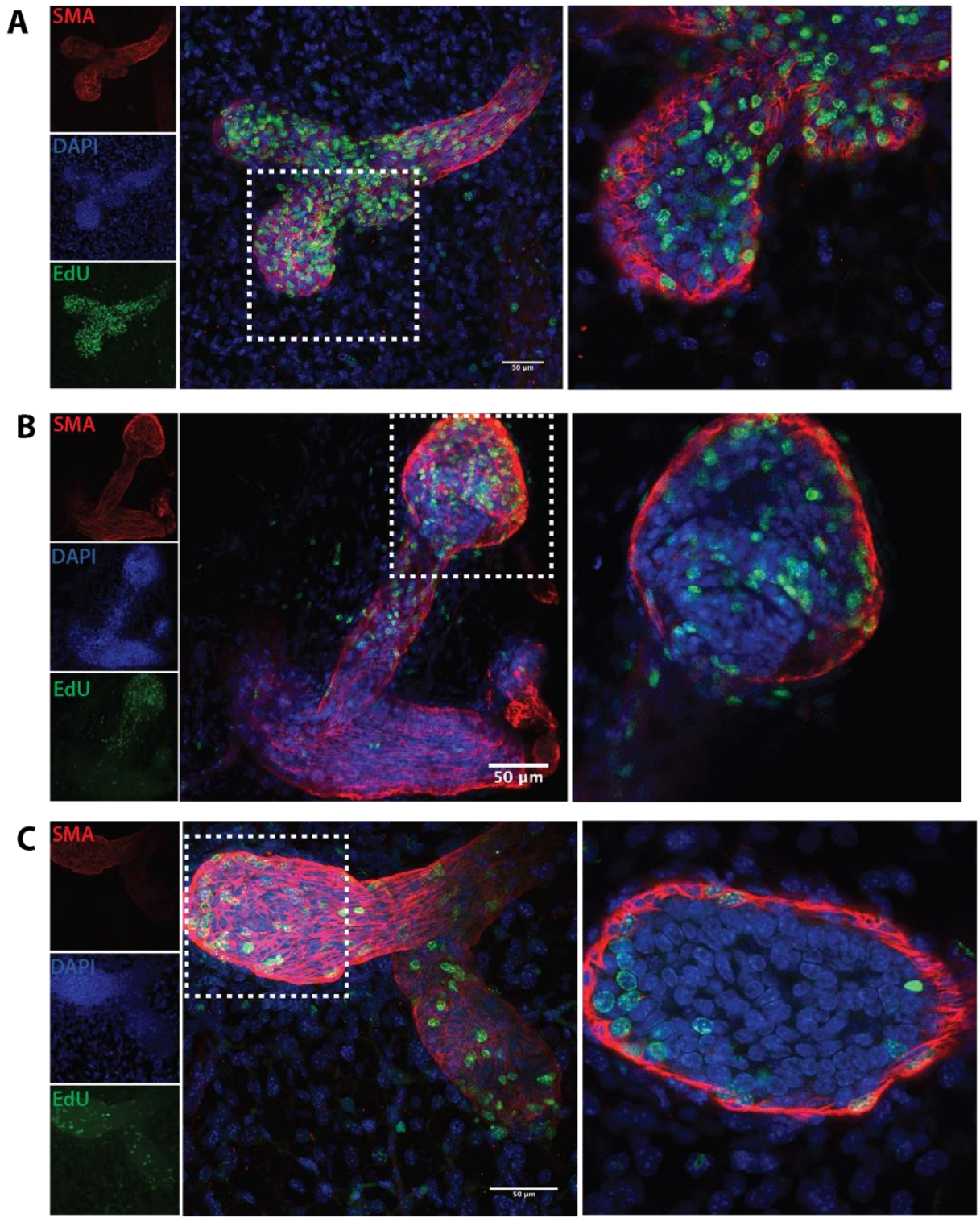


Figure 5.2.18 EdU proliferation seen in TEBs of 7 week mammary glands

Proliferation seen in TEBs as measured by EdU incorporation in 7 week-old-mice. Proliferation can be seen in both cap and body cells, and also in the subtending ducts. Representative images from mammary glands of n=10 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

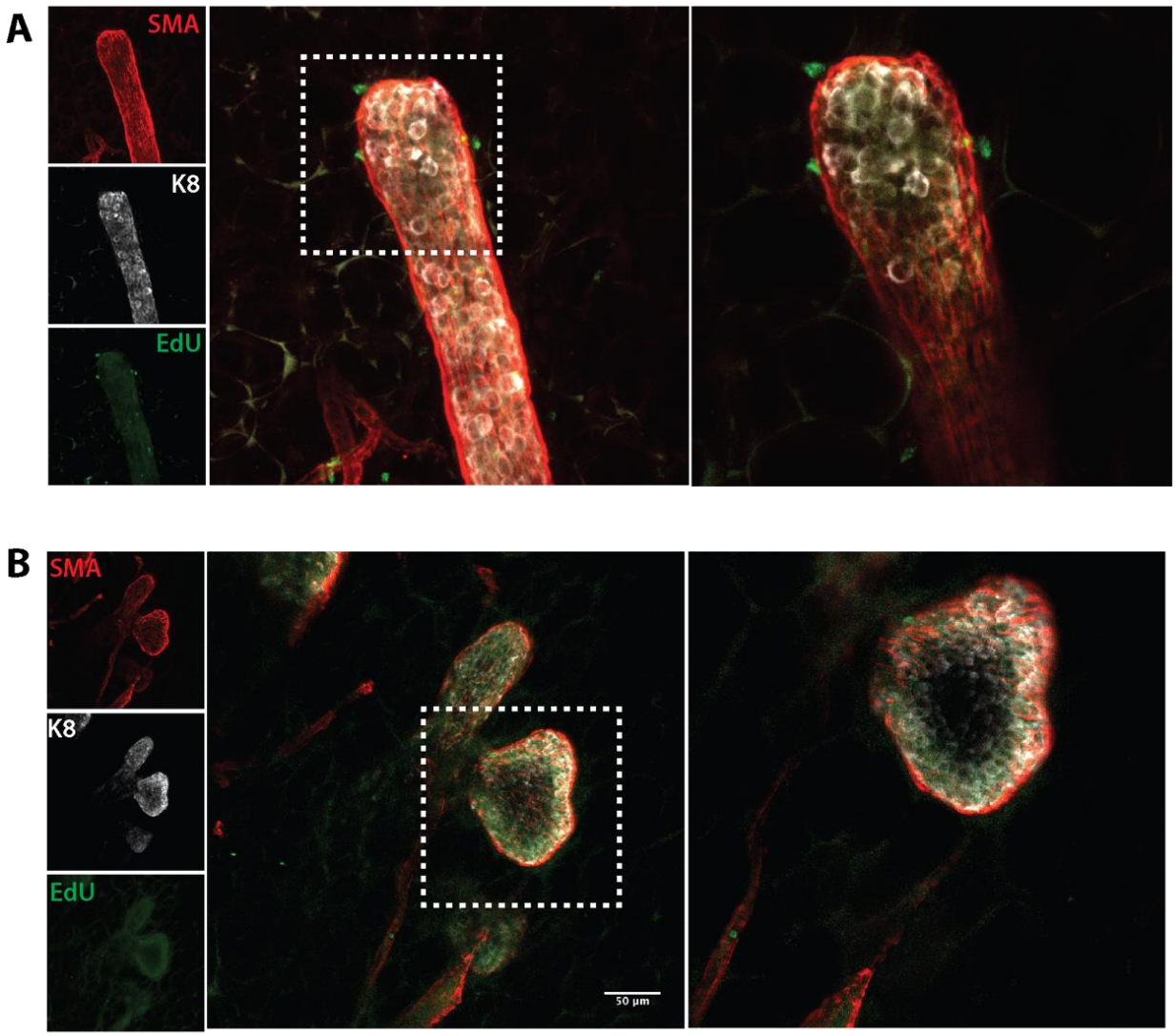


Figure 5.2.19 Absence of proliferation in 7 week mammary glands

Representative images of regions in later pubertal mammary glands, showing that not all regions remain proliferative. These include both terminal end ducts **(A)** and side branches **(B)**. Representative images from mammary glands of n=10 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of regions of interest (boxed and enlarged).

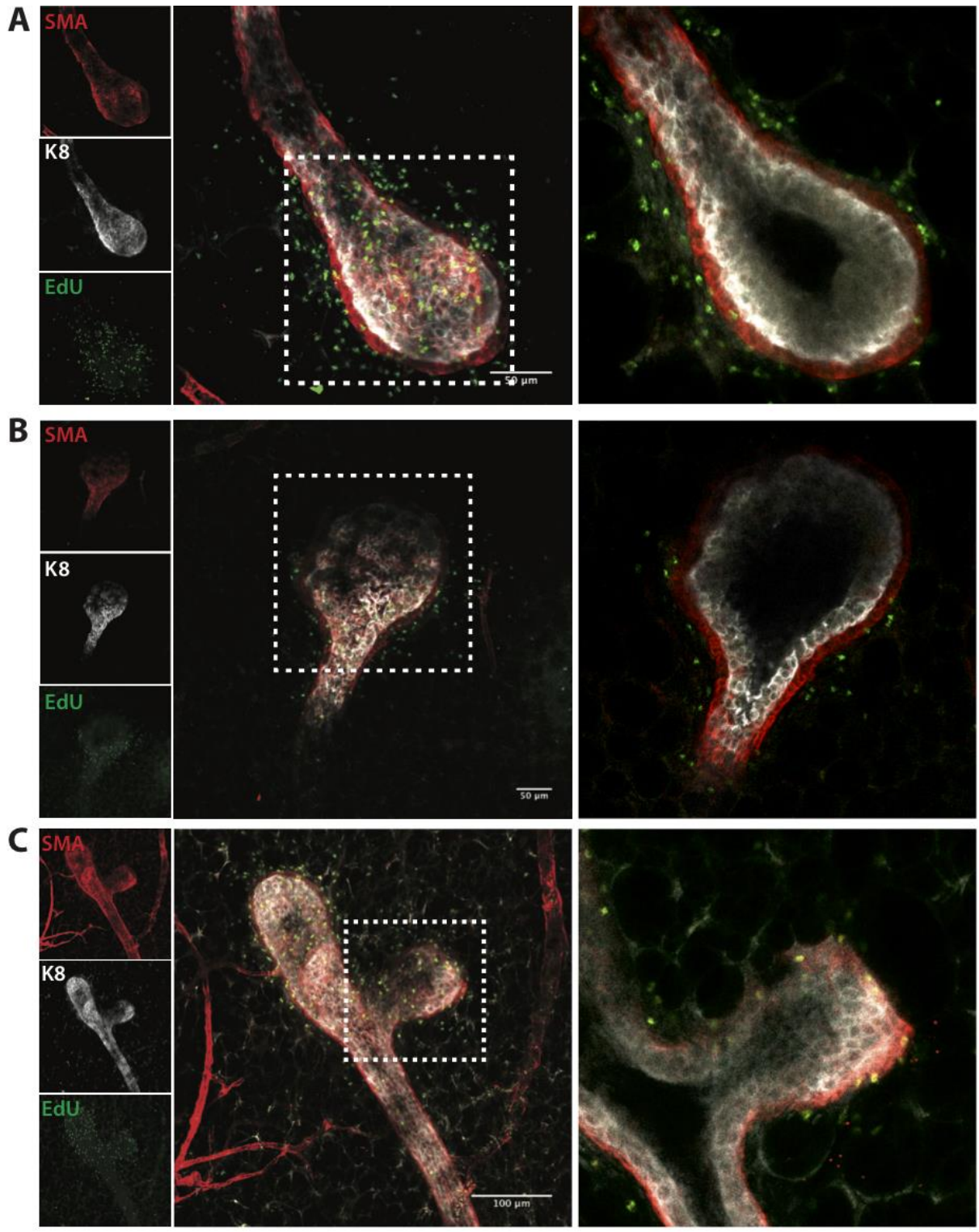


Figure 5.2.20 Proliferation seen in stromal cells in 7 week mammary glands

Images show examples of proliferation seen in stromal cells surrounding TEBs during the latter phase of pubertal development. This is both in TEBs (A,B) and in bifurcating regions **(C)** Representative images from mammary glands of n=10 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

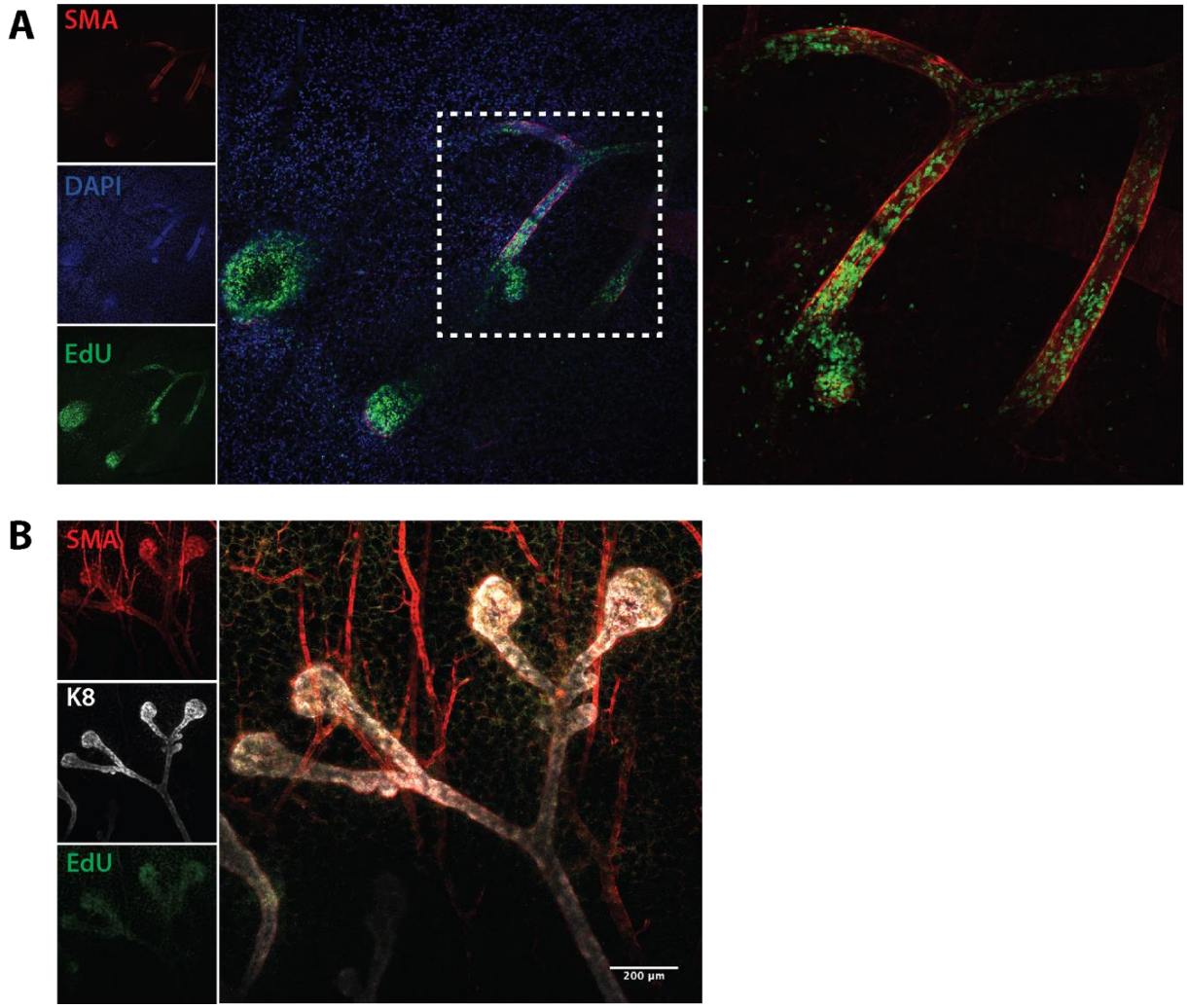


Figure 5.2.21 Proliferation in ductal regions of 7 week mammary glands seen by EdU incorporation

(A) Representative image showing EdU incorporation in both TEBs, and in the subtending ducts **(B)** Representative image demonstrating absence of proliferation in both TEBs and ducts. Data from mammary glands of n=10 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

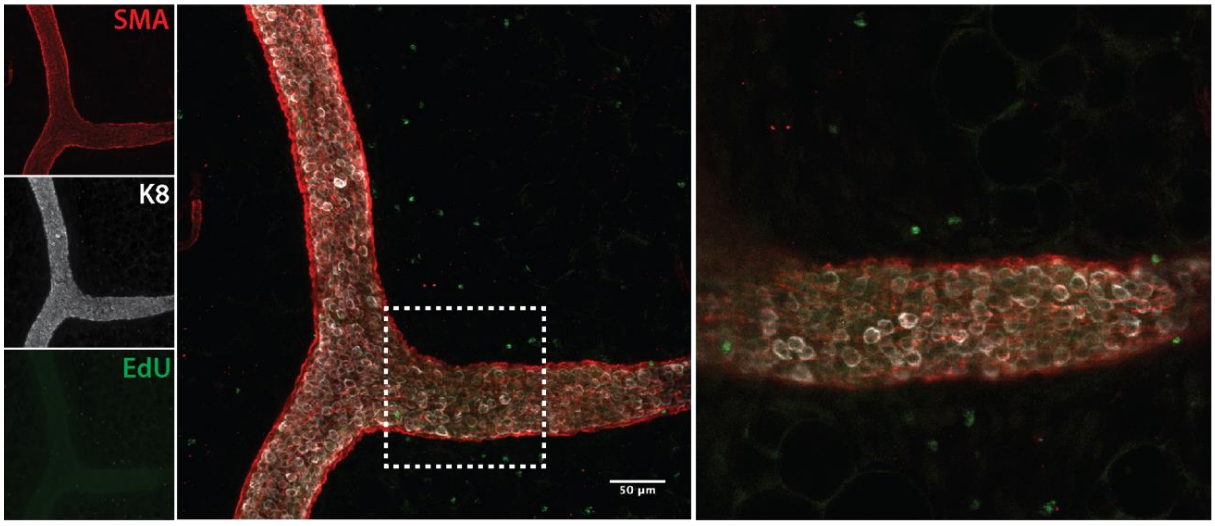


Figure 5.2.22 Absence of proliferation in ductal regions of 7 week mammary glands

Representative image demonstrating absence of proliferation, shown by no EdU signal, in a developed ductal region in 7 week mammary gland. Data representative from mammary glands of n=10 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

5.3.1.1 Post-pubertal EdU proliferation

After pubertal morphogenesis has finished, proliferation in the gland returns to a much lower level, fulfilling only expansion during oestrous cycling. At 12 weeks, pubertal hormones will have diminished and the gland will have undergone a number of rounds of cyclical proliferation and regression with oestrous cycling. This could be seen with increased side branching of the ducts in 12 week-old animals. However, despite this proliferation levels were very low with only sparse EdU+ cells identified, both at the terminus of ducts (Figure 5.2.23a) and to a lesser extent within the ductal network (Figure 5.2.23b).

In a similar vein, the mammary glands of 24 week animals were much more complex, with both increased side branching and vasculature (Figure 5.2.24a). EdU incorporation at this stage was also low, with the majority of ducts showing no proliferative activity (Figure 5.2.24b). A few, sparse labelling events could be detected, which were largely found within the secondary side buds and branches (Figure 5.2.25a,b). These presumably represented cells which have proliferated on hormonal stimulation with oestrous cycling. This reduction in proliferation compared with younger mice is to be expected due to the diminished requirement for such spectacular epithelial expansion and ductal elongation.

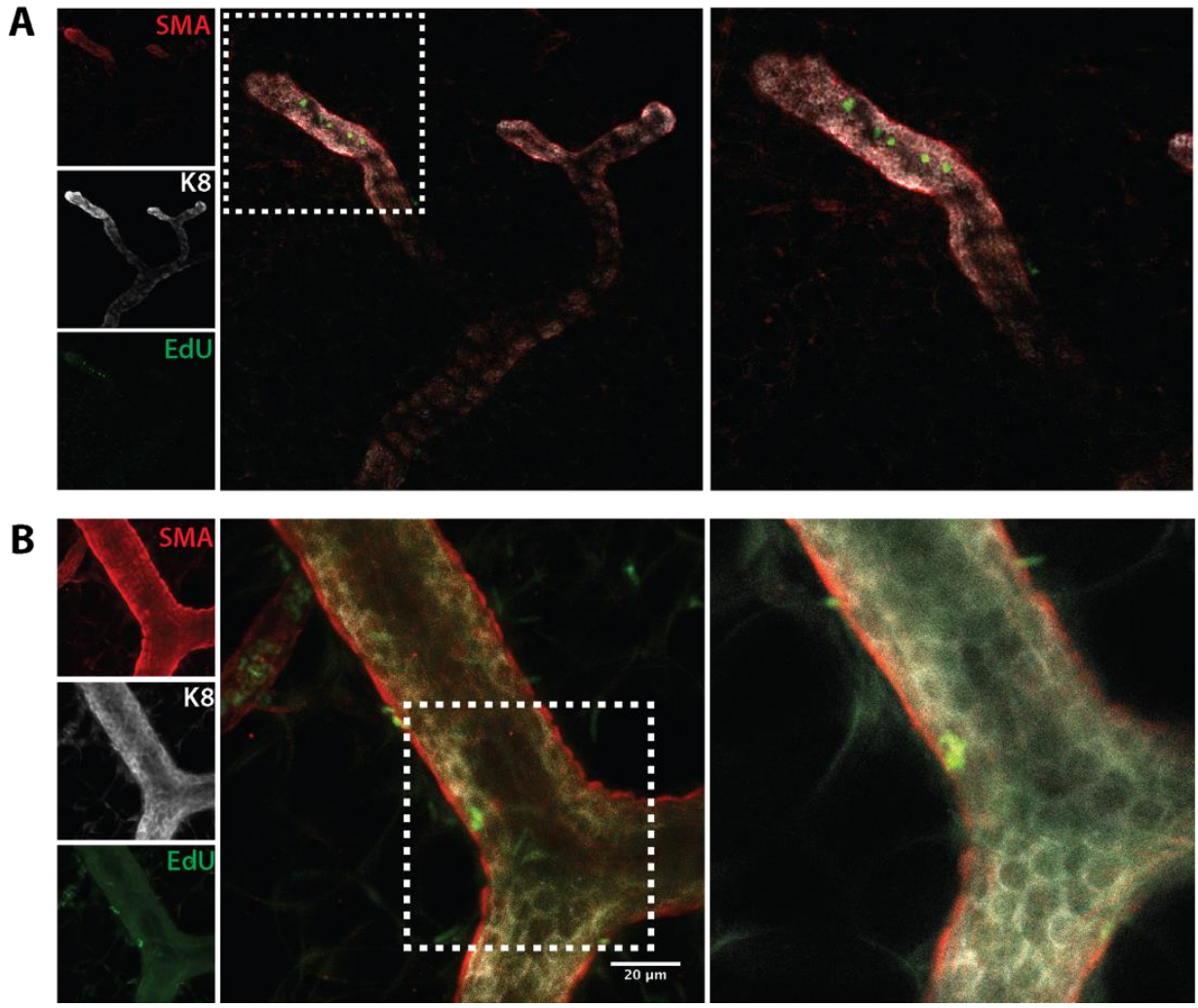


Figure 5.2.23 Proliferation in 12 week adult mammary glands seen by EdU incorporation

Examples of EdU incorporation seen in 12 week mammary glands. **(A)** Representative image showing sparse labelling in terminal end duct. **(B)** Example of a rare EdU+ cell seen within the ductal region. Images are from mammary glands of n=10 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

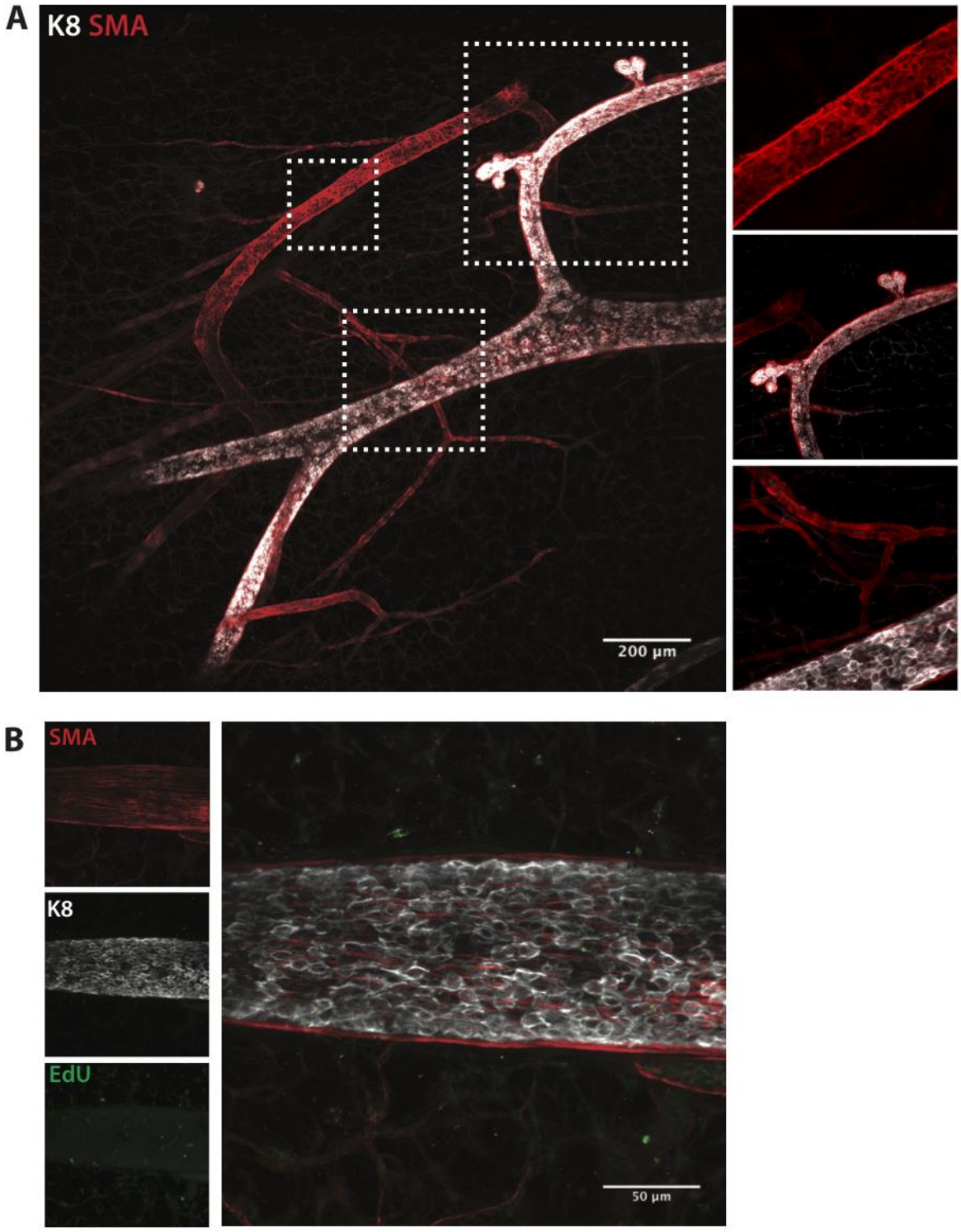


Figure 5.2.24 EdU incorporation investiageted in mammary glands of 24 week mice

(A) Representative image demonstrating the increased complexity in both the ductal and vascular network of older adult mammary glands. **(B)** Representative image showing the absence of EdU incorporation. Maximum-intensity z-projections of

individual channels, merge and optical slices of a regions of interest (boxed and enlarged) from mammary glands of n=10 mice.

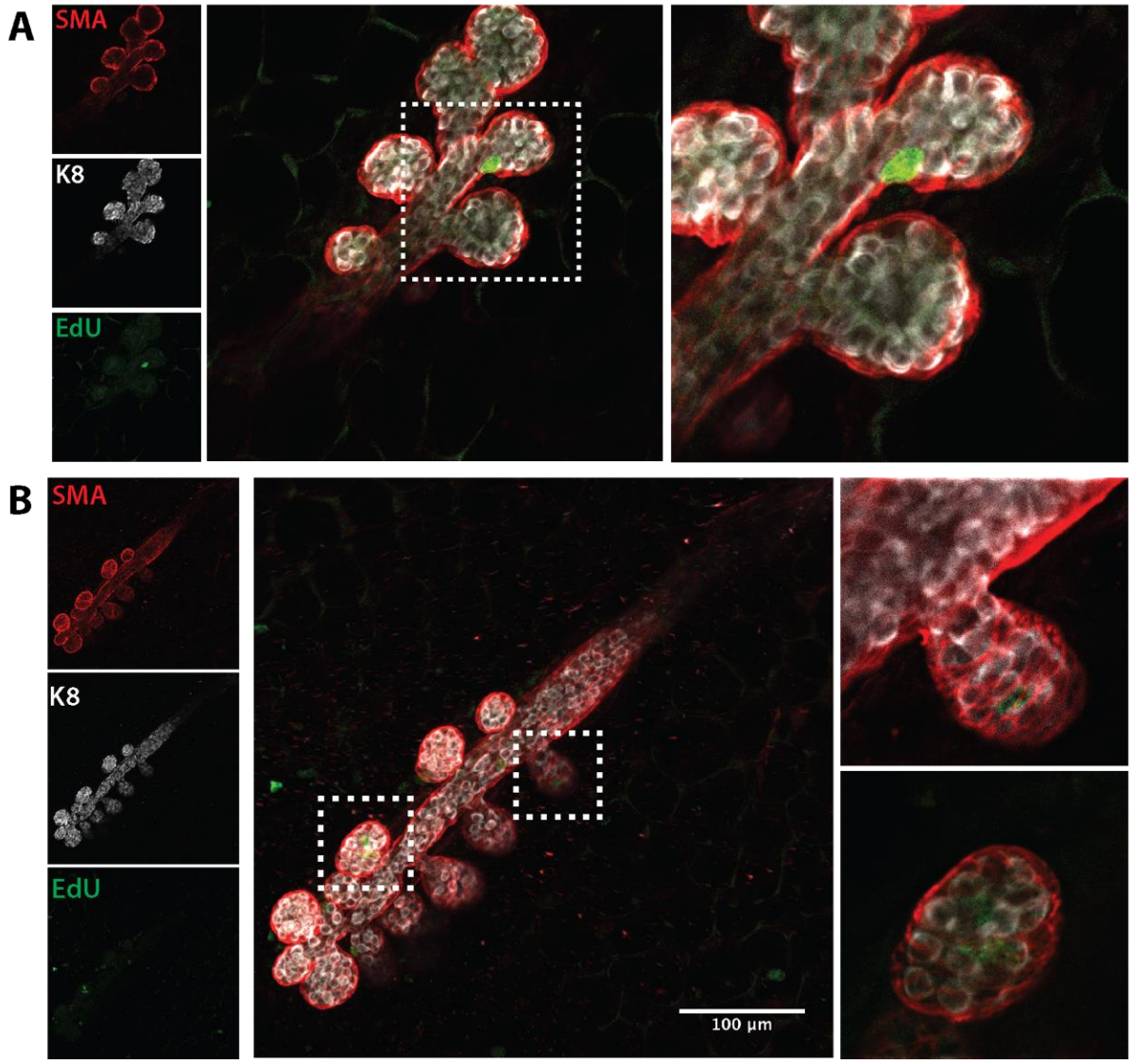


Figure 5.2.25 Proliferation in 24 week mammary glands using EdU incorporation

Examples of EdU incorporation seen in mammary glands of 24 week mice, seen in side branches and buds. **(A)** and **(B)** show representative images from mammary glands of n=10 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a regions of interest (boxed and enlarged).

5.2.4 Use of the Fucci2a system to study mammary gland proliferation

Genetically engineered transgenic mouse models have also been developed to allow visualisation of the cell cycle without the need for DNA intercalators, which also provide a higher resolution of cell cycle stage. Indeed, one such model, the Fluorescence Ubiquitin Cell Cycle Indicator 2a (Fucci2a) mouse, allows for a more distinct separation of the cell cycle stage by utilising mCherry tagged Cdt1 and mVenus tagged geminin (Mort et al., 2014). These proteins are expressed at distinct phases of the cell cycle and are then rapidly degraded. The expression of these reporters thus results in the ability to visualise cells in different stages of the cell cycle: cells in G1/0 express just Cdt1 as therefore mCherry alone, cells in G1/S transition express both Cdt1-mCherry and geminin-mVenus so can be seen as yellow, and those cells in late S/G2/M will be green due to geminin-mVenus expression. Furthermore, cells that have recently divided and so have not yet accumulated enough Cdt1-mCherry will show no reporter expression and can be distinguished as such (Figure 5.2.26a). Importantly, this is driven by the ubiquitous R26 promoter and so the construct should be expressed in all cells.

Accordingly, to validate the results seen with EdU and Ki67, the mammary glands from 12 week virgin mice (n=12 mammary glands from 3 mice) were investigated for cell cycle stage. Tissues were stained with the luminal cell marker K8 and optically cleared using SeeDB optical clearing to ensure the reporter fluorescent integrity was not affected. In consensus with the data from EdU incorporation, the majority of cells visualised showed expression of Cdt1-mCherry, suggesting that the cells were in either G1 or G0 (Figure 5.2.26b). Given these mice were virgins, it is likely that the majority of Cdt1 expression in these cells is associated with the G0 stage of the cell cycle.

However, a number of cells also showed no reporter expression – suggesting a recent cell division (Figure 5.2.26b, arrow heads). However, it may also be that this is a consequence of the 3D imaging meaning reporter detection is affected, given the mature gland sizes. Furthermore, issues with transgene infidelity have been reported resulting in lack of expression in some cells (Mort et al., 2014). This is highlighted here, shown by the surrounding stromal cells, in which no expression of either reporter was seen in any cells (Figure 5.2.26b,c) – again, a phenomenon that has been reported in other systems.

There were also rare instances of cells co-expressing both mVenus and mCherry, suggesting these cells were in G1/S transition. These cells were seen both in the tips of ducts (Figure

5.2.27a), in addition to further down the ductal tree (Figure 5.2.27b). It is tempting to speculate that these cells may represent those that are proliferating upon hormonal stimulation with oestrous cycling. Rare cells in the S/G2/M phases of the cell cycle, indicated by mVenus-Geminin expression, were also detected, again largely at the tips of ductal regions (Figure 5.2.28a,b,c).

Together, these data highlight the utility of the Fucci2a system in the mammary gland to study proliferation, combined with 3D imaging. Further work examining this at different ages, utilising FACS to measure mVenus and mCherry expression would provide valuable quantification of cell cycle states, as has been recently shown in different oestrous cycle stages (Shehata et al., 2018). Importantly, these data corroborate that seen with both EdU incorporation studies and Ki67 staining, in that post-pubertal mammary glands have greatly reduced proliferation in comparison to younger animals

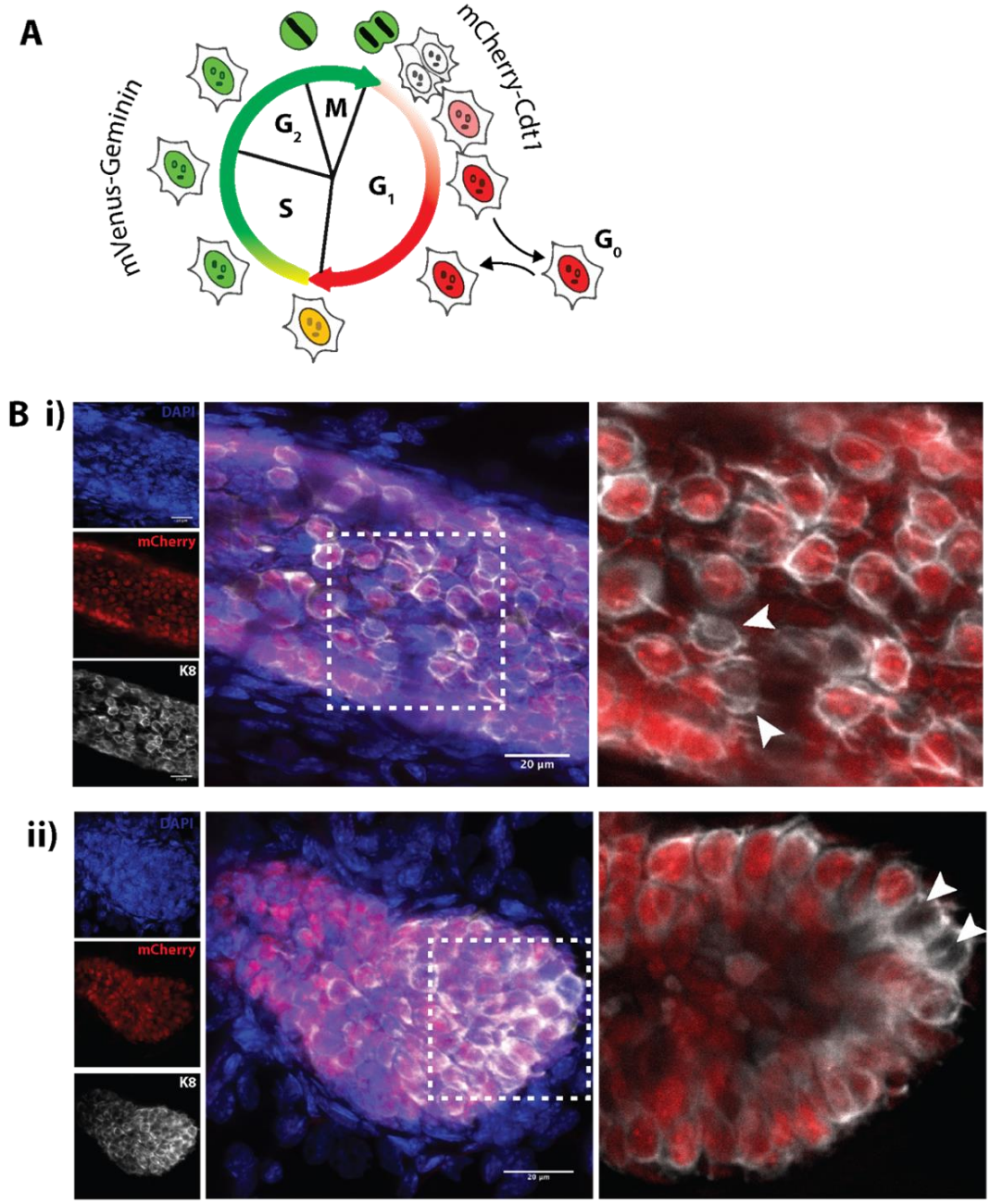


Figure 5.2.26 Use of Fucci2a systm to study cell cycle in 12 week mammary glands

(A) Schematic of the Fucci2a system, which uses mCherry tagged Cdt1 and mVenus tagged Geminin to visualise cell cycle stages. Cells in G1/G0 will be mCherry+ (red), whilst cells in G1/S transition will be mCherry+ and mVenus+ (yellow) due to co-expression of both reporters. Cells in S/G2/M will be mVenus+ (green), with cells just divided having no reporter expression **(B)** Examples of Fucci2a transgenic mammary glands from 12 week-old mice co-stained with K8 for luminal cells. The majority of cells express mCherry indicating G1/G0 phase, however some cells were also seen to have no reporter expression (arrow heads), suggesting they have just divided. Images are representative from mammary glands of n=3 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

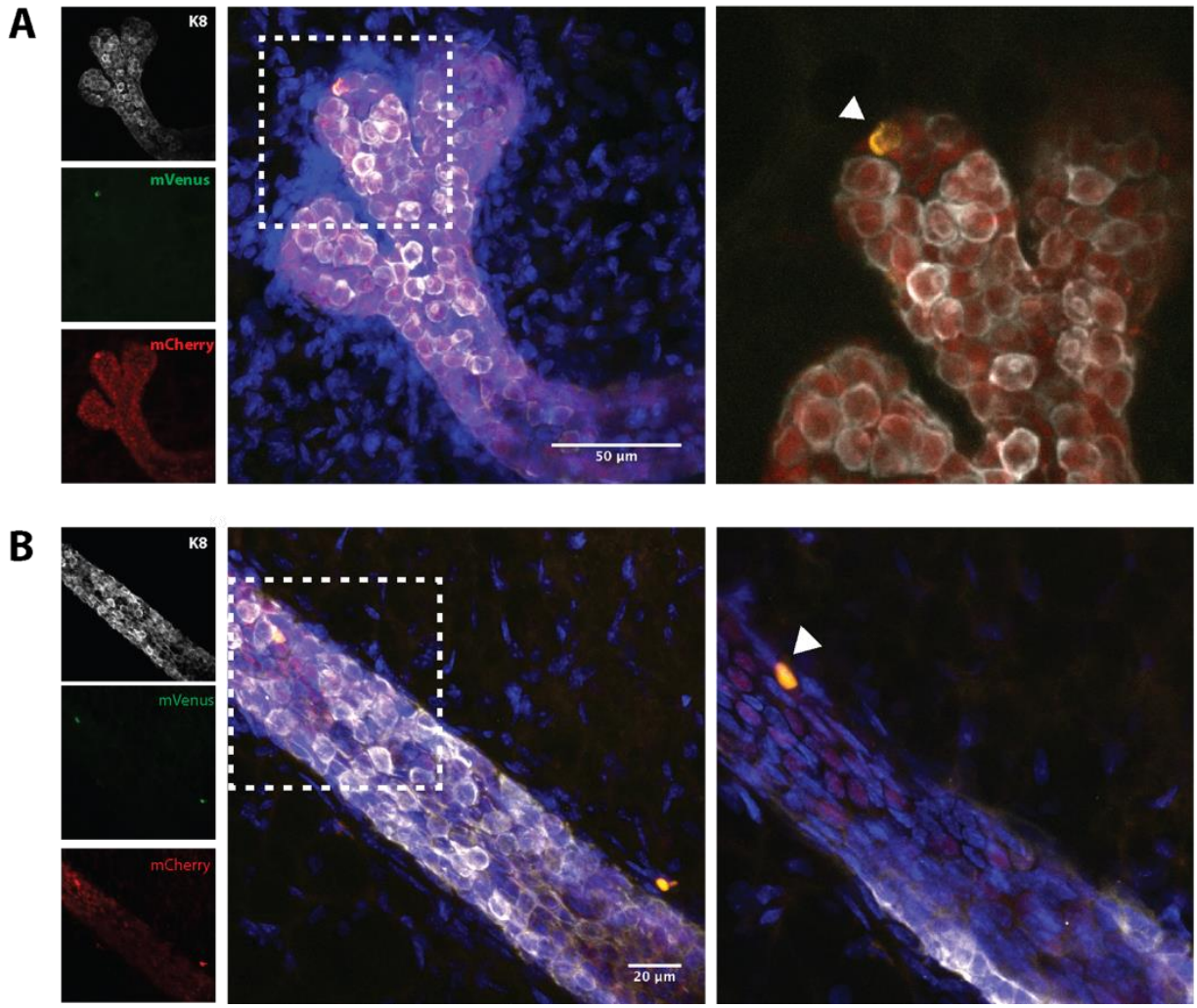


Figure 5.2.27 Fucci2a system in 12 week mammary glands shows cells in G1/S transition

Examples of Fucci2a transgenic mammary glands from 12 week-old mice co-stained with K8 for luminal cells and DAPI. **(A)** and **(B)** show representative images of rare yellow cells, expressing both reporters mVenus-Geminin and mCherry-Cdt1, suggesting these cells are in G1/S transition (arrowhead). Images are representative from mammary glands of n=3 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged).

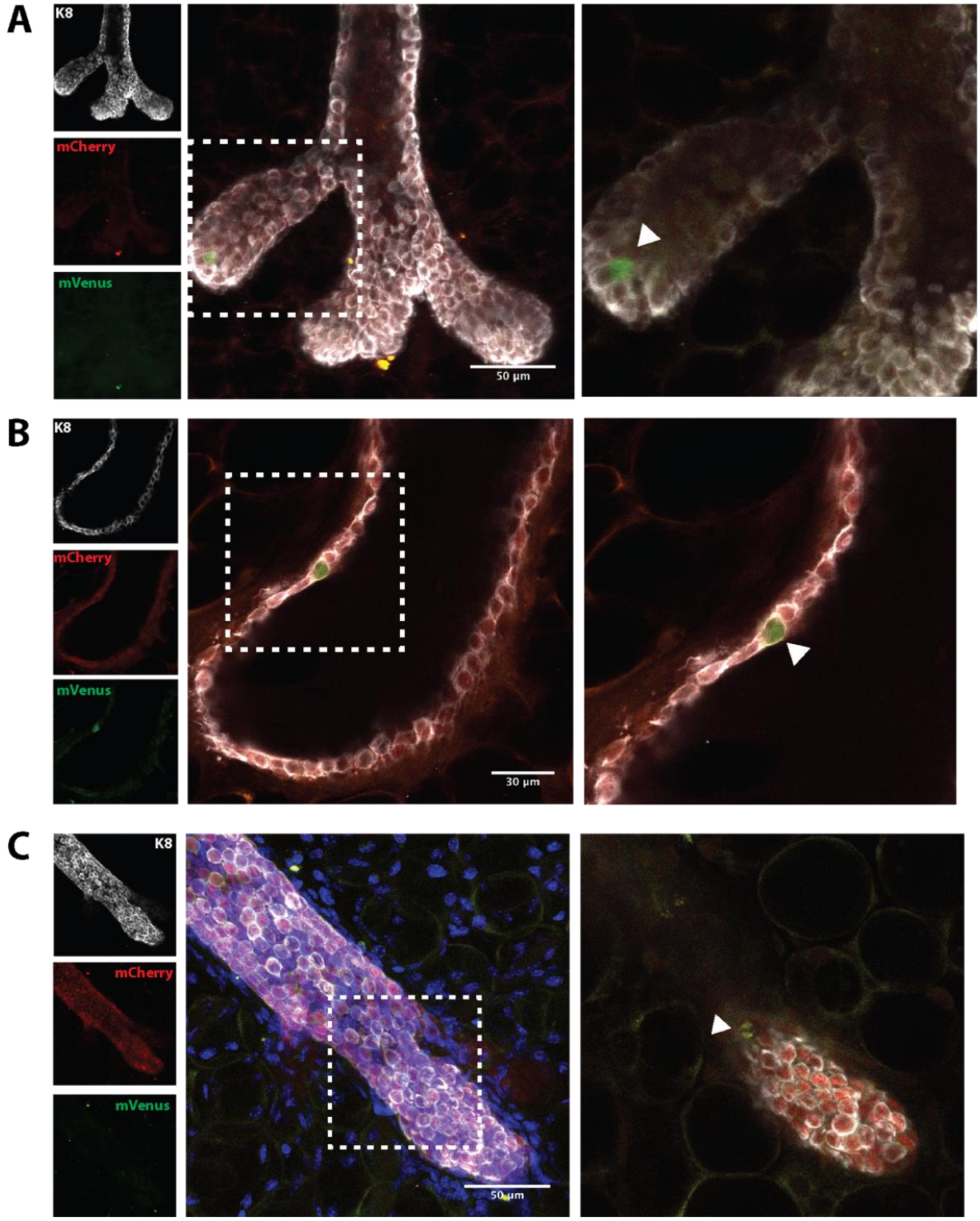


Figure 5.2.28 Fucci2a system in 12 week mammary glands shows cells in S/G2/M transition

Examples of Fucci2a transgenic mammary glands from 12 week-old mice co-stained with K8 for luminal cells and DAPI. Representative images of rare green cells with mVenus-Geminin expression, suggesting these cells are in S/G2/M phase of the cell cycle. Images are representative from mammary glands of n=3 mice, showing maximum-intensity z-projections of individual channels, merge and optical slices of a region of interest (boxed and enlarged.

5.3 Discussion

By examining proliferation in a variety of ways, combined with optical tissue clearing and 3D imaging, these data have provided fascinating insights into the development and homeostasis of the mammary gland. Using wholemount imaging, we have also shown where the proliferative regions in the mammary gland are located, and that levels dramatically reduce with age. By combining EdU incorporation, Ki67 staining and the Fucci2a system, these data have provided a useful tool to assess proliferation and the cycling nature of functional, proliferative progenitor/MaSCs when overlaid with the lineage tracing models.

5.3.1 Overview of main findings

One of the most interesting aspects of these data relates to the finding that proliferation within the pubertal mammary gland is not restricted to the TEB. Indeed, the result that proliferation could be detected in the duct is an unexpected result that has not been previously described; proliferation and growth in the developing mammary gland is usually reported to be mediated by the TEBs (Kenney et al., 2001; Hinck and Silberstein, 2005). However, this information is generally obtained from 2D sections, and as such this can be difficult to infer how that data fits with the 3D topology of the gland. Here, 3D imaging has allowed areas to be imaged in the context of the entire ductal network. This was first seen with Ki67 staining but given this marks cells in any stage of the cell cycle, combined with the rapid growth of the gland during pubertal development, this does not provide definitive information. However, the subsequent use of EdU incorporation and a relatively short chase period demonstrates that these cells detected in the ducts are in fact actively proliferating.

Several recent studies have reinforced the traditional view that the TEB is responsible for ductal elongation, even when using 3D or intravital imaging (Scheele et al., 2017). This highlights the potential limitations of these imaging methods (outlined in chapter 3) – or in the case of intravital imaging, the challenges associated with imaging at a high enough resolution. Given the notion that stem cells and progenitors are found throughout the ductal network, supported by both early transplantation studies and fate mapping work carried out the previous chapter of this thesis, it is logical that stem/progenitor cells that are deposited by the travelling TEB are able to contribute to the elongation of the gland, in addition to secondary branching of the duct.

It may also be that these ductal regions represent areas where functional stem or progenitor cells are found, which could represent a "transit-amplifying" population, acting as an intermediate population able to support the regenerative capacity of the gland. Indeed, this has been established in other epithelial tissues, including both relatively more quiescent systems such as the hair follicle (Hsu et al., 2014) and highly regenerative tissues such as the crypts of the intestine (Snippert et al., 2010). However, this raises important questions surrounding the signals involved in this. Whether this proliferation is cell-intrinsic or in fact these cells have been called to proliferation by stromal or niche signals – and whether these signals may be hormonal (for example oestrogen or progesterone) or rather signalling factors released by local stromal cells such as macrophages remains to be determined.

Traditionally, it has been suggested that MaSCs reside in the basal population, and as such it would be expected that it would be these cells that proliferate during ductal morphogenesis (Plaks et al., 2013; Rios et al., 2014; Wang et al., 2015). However here, proliferation was detected in both the luminal and basal lineages during this time. Further in-depth quantitative analysis would be useful to illuminate definitively the proportion of luminal and basal EdU+ cells at different stages of development and within different regions of the ductal system. However, this finding also aligns with the conclusions of the previous chapter, in that two different lineage restricted stem/progenitor cell populations produce cells of the luminal and basal lineages.

As discussed in the introduction, the cycling nature of MaSCs is not yet fully understood. Consequently, it is important to appreciate that the approaches utilised here would not detect the existence of quiescent stem or progenitor cells within the mammary gland. Indeed, these cells would not be labelled by EdU or Ki67, nor can they be distinguished from cells in G1 with the Fucci2a system. For this, long term label retention studies would be required using either EdU or H2B-GFP to detect cells which have remained over time. These have been previously used (Smith, 2005; Boras-Granic et al., 2014), but combining them with 3D imaging would provide further resolution of the nature of these cells. Alternatively, newer Fucci models have been developed, including one system (termed FUCCI4) which allows for the distinction between all phases of the cell cycle by utilising four different fluorescent indicators, as opposed the two used in the Fucci2a system here (Bajar et al., 2016).

Using both Ki67 staining and EdU incorporation, high levels of stromal cell proliferation were detected. Unfortunately, spectral limitations prevented co-staining of stromal cell markers in

addition to DAPI, EdU and labelling of basal and luminal cells. However, it would be interesting to determine the identity of these proliferating stromal cells seen both in early and later puberty. Indeed, a number of immune cells (including mast cells, eosinophils and macrophages) have been implicated in supporting ductal elongation in puberty, but the degree to which these cells proliferate and expand *in situ* within the gland, or home from elsewhere has yet to be determined.

Further, the degree to which the lymph nodes, which are found in close proximity to the mammary glands, may provide this stromal support is a matter that requires further investigation. This is a particularly interesting aspect, given the frequency with which lymph nodes are the primary sites of metastasis in breast cancer (Li et al., 2014). Moreover, there is increasing understanding of the relevance of the immune system in a number of malignancies, including breast cancer (Azizi et al., 2018), and so understanding the normal immune environment is vital.

5.3.2 Oestrous cycling relevance

An important consideration when studying proliferation in the postnatal mammary gland is normalising for the oestrous stage of the mice. Indeed, particular attention to this aspect was paid here given the heterogeneity that has been reported. Indeed, this is highlighted by previous work which showed that the numbers of basal and MaSCs (indicated by mammary repopulating units, MRU) can fluctuate 4 and 14-fold respectively in FVB mice, depending on the stage of oestrous (Giraddi et al., 2015). By using MRUs as a measurement of stem cell activity, it has also been previously observed that the number of stem/progenitor cells are at a maximum during dioestrous (Joshi et al., 2010).

Furthermore, in C57BL/6J mice it has been suggested that basal cells increase 3-fold in dioestrous compared with other oestrus stages, in addition to an up to 7-fold increase in MRUs. Finally, it has also been shown that cells in dioestrous are especially sensitive to the S-phase cytotoxin 5FU, with a 75% reduction in MRU number upon exposure (Giraddi et al., 2015). Together, these data provided the impetus to examine mice in dioestrous, under the assumption that this is the most active stage of the oestrous cycle.

However, more recent work has highlighted the heterogeneity even within this stage of the oestrous cycle. Utilising the Fucci2 mouse model together with sequential administration of DNA intercalators, these authors further highlighted the heterogeneity in both the cell cycle lengths of mammary epithelial cells, and the length of oestrous stages. As has been previously suggested,

they also demonstrated that there appears to be two distinct states of dioestrous, whereby cells are either highly proliferative or not. Furthermore, the inter-cycle heterogeneity was shown using CFSE labelling, which demonstrated that the proliferative expansion associated with dioestrous does not necessarily occur each cycle (Shehata et al., 2018). Here, we attempted to control confounding factors such as this by oestrous staging the mice and both injecting and culling at the same time every day. Furthermore, mice were examined for correct cycling for 7 days prior. However, despite controlling for these factors, evidently it is not possible to be entirely certain when the mice cycle through the different stages, or to ensure that they do so at the same rate.

5.3.3 Importance and implications

Identification of proliferative cells and regions in the mammary gland has implications for breast cancer. If cell populations are able to survive multiple cycles of proliferation and death throughout both oestrus cycles and multiple pregnancies, it raises the possibility that with the addition of oncogenic hits or pro-survival signalling from the niche then these could be the cells of origin, or at least contribute, to the development and maintenance of breast cancer.

Previous work by others using gene expression profiling has shown that the genetic profiles of subpopulations of luminal cells in both mice and humans resemble that of two subtypes of breast cancer, luminal A and luminal B (Lim et al., 2010; Shehata et al., 2012), which account for 50% of breast cancers. Moreover, Ki67 has been implicated multiple times as a prognostic marker in breast cancer (Inwald et al., 2013), demonstrating the importance of understanding proliferation in the normal mammary gland

5.3.4 Overall conclusions

Overall, the results presented in the chapter highlight the vast proliferative capacity of the pubertal mammary gland, and highlights the importance of using multiple methods combined with 3D imaging. Moreover, the finding of ductal proliferation during pubertal development provides support to the hypothesis that MaSCs may be laid down along the duct from the TEBs as the gland develops.

Chapter 6

Assessment of an *ex vivo* organoid model to study mammary gland development

6.1 Introduction

The work presented in this thesis has thus far have aimed to examine aspects of mammary gland development using agnostic and unbiased methodologies *in vivo*. Through the use of such techniques, we have highlighted a number of novel insights into mammary gland and stem cell biology. In this chapter, we sought to determine whether our observations were recapitulated in an *ex vivo* organoid culture model. Although potentially less faithful to true biological processes than their *in vivo* counterparts, *ex vivo* culture systems offer numerous opportunities for rapid and complex perturbation studies, for example CRISPR/Cas9 gene editing or small molecule drug treatment, which can aid in a holistic understanding of mammary gland development in health and disease.

6.1.1 Background of 3D culture models in the mammary gland

Although *in vivo* models are generally thought to have higher fidelity to processes of human health and disease, *ex vivo* culture systems have specific utility in perturbation and high throughput studies. For example, a recent study by Bruna *et al*., established a biobank of more than 80 patient derived xenograft (PDX) models of breast cancer (Bruna et al., 2016). The authors found that although the models were more representative of their originating samples when expanded *in vivo*, it was only when *ex vivo* culture systems were established that high throughput drug screens were possible.

Over the last five decades, *in vitro* and *ex vivo* culture methodologies for human mammary cells have become more complex, drawing on our increasing understanding of the complexity of signalling in the tissue niche. Indeed, it has been demonstrated that a number of cell signalling networks are significantly altered when cells are in 2D rather than 3D – highlighted by the fact that drug screening outcomes often are not replicated when taken into an *in vivo* setting. Even when controlling for a lack of cellular diversity, factors often overlooked in *in vitro* culture systems such as cellular tension and surface nanotopography can have profound effects on cellular behaviour (Cassidy, 2014). Traditionally, *in vitro* cultures of mammary cells (for example MCF10As) have largely comprised monolayers cultured on tissue culture plastic. The first 3D culture model of the mammary gland was described in 1977, utilising floating collagen membranes in a medium containing insulin, hydrocortisone and prolactin. This system was composed of a continuous "pavement" of mammary epithelial cells with some alveolus-like structures that could be maintained for one month in culture (Emerman and Pitelka, 1977). Since this time, more complex collagen-based cultures have been developed, utilising both scaffolds (Campbell et al.,

2017; Hume et al., 2018b) and gels to study aspects of mammary biology, including branching morphogenesis and migration of mammary cancer cells.

Recently, the development of ECM-based gels has led to an explosion in protocols for so-called "organoid" systems, largely pioneered by culture of single Lgr5+ intestinal stem cells into mini crypt-villus structures (Sato et al., 2009). Since then, a number of other organoid culture systems have been developed for other epithelial tissues, including the liver (Huch et al., 2014), stomach (Barker et al., 2010), and pancreas (Huch et al., 2013a). This has led to confusion in the nomenclature, with the term organoid having different definitions depending on the tissue of origin. Largely, organoids can be defined as *in vitro* or *ex vivo* cultures of 3D cellular clusters that have been derived from primary tissue, embryonic stem cells, or induced pluripotent stem cells, which are capable of self-renewal and self-organisation, and can exhibit similar organ functionality as the tissue of origin (Fatehullah et al., 2016).

Given the uncertainty of the identity of the stem cells within the mammary gland, the term largely refers to primary explants of epithelial ducts into 3D ECM gels, as opposed to clonal cultures derived from single cells (Shamir and Ewald, 2014). A number of ECM gels have been developed for this, but Matrigel is the most widely used, having been originally described in 1977. It is derived from the secretions from tumours of the Engelbreth-Holm-Swarm (EHS) mouse sarcoma cell line (Orkin et al., 1977), with its major constituents being collagen IV, laminin and fibronectin, in addition to a number of growth factors including TGFβ and EGF – all of which are components of the basement membrane in the mammary gland

Consequently, a number of organoid culture systems for the murine mammary gland have recently been described which have contributed significantly to our understanding of the biology of the mammary gland, providing insights into aspects such as differentiation, hormonal regulation, paracrine signalling and cell-cell interactions (Mroue and Bissell, 2013; Shamir and Ewald, 2014). Some of these methods have shown that cultures can be derived from single, FACS sorted mammary epithelial cells (Jamieson et al., 2017). However, the process of dissociation the mammary gland into single cells for this process has been shown to disrupt the normal biology. Moreover, this prevents inclusion of the mammary mesenchyme, including stromal cells, which have been shown to be vital in supporting mammary gland development and homeostasis.

A protocol that does not require dissociation into single cells for downstream FACS and allows for long term culture that recapitulates the normal mammary gland was chosen. This method utilises a medium containing growth factors which have been shown to be important for mammary gland development and maintenance including Neuregulin1 (Nrg1) and R-spondin 1 (Rspo1) (Cai et al., 2014). Importantly, vital aspects of mammary gland biology are maintained, including luminal/basal cellular organisation and hormone receptor expression in luminal cells (Jardé et al., 2016). Indeed, organoids were even seen to express milk protein when stimulated with prolactin to induce differentiation, demonstrating that they retain this important functionality.

6.1.2 Importance

The utility of *ex vivo* models which recapitulate the tissues of origin have far reaching uses. Indeed, well characterised organoid culture systems have been developed for a number of other epithelial tissues as mentioned above, which have been used in the study of both homeostasis and carcinogenesis. For example, a biobank of 20 human colorectal carcinoma (CRC) organoids has recently been described that largely recapitulates molecular features of the originating tumour sample (van de Wetering et al., 2015). Through exome-sequencing and RNA expression analysis, the authors were able to show that the CRC organoid biobank captures most of the mutational and expression landscapes observed in large CRC studies. Furthermore, screening of these cultures was carried out using an 83-compound library to identify molecular signatures associated with drug responses, highlighting the utility of *in vitro* 3D cultures in the drug discovery process.

The development of an organoid culture system in the mammary gland would provide a model to study developmental mechanisms and tease out specific pathways of interest, in addition to showing translational benefit in breast cancer drug screening studies. In order to develop such a culture methodology, we sought to qualitatively assess the fidelity of organoid culture protocols for maintaining features of mammary gland biology, and mammary stem cell behaviour, elucidated and described earlier in this thesis.

6.1.3 Aims

This chapter aims to assess the use of 3D organoid culture to study aspects of mammary gland development that have been explored in this thesis, mainly assessed through the use of the three techniques described in the previous chapters: 3D imaging using optical clearing, lineage tracing using the $R26^{[CA]30}$ model and examining proliferation with EdU incorporation.

6.2 Results

6.2.1 Assessment of organoid biology

As discussed, a number of methods have been developed for 3D *ex vivo* culture of the mammary gland. In the system used here, mammary epithelial cells are derived from the glands of adult virgin mice. Briefly, tissues were mechanically minced then enzymatically digested before a brief fibroblast depletion step to remove the bulk of the fibroblasts. After this, cells were re-suspended in Matrigel, plated and maintained in culture with organoid media for up to four months. After only a day in culture, small spherical organoids could be observed (Figure 6.2.1,i), which were seen to begin to branch in under a week (Figure 6.2.1,v). We next sought to investigate the organoids in more detail, focussing on how well they recapitulated the biology of the normal mammary gland. As such, we moved to higher resolution imaging using confocal microscopy.

Due to the 3D nature of the organoids, it is important to be able to observe them in multiple focal planes to fully appreciate their morphology, similar to the *in vivo* mammary gland. Some of the older cultures were seen to be particularly large, at times reaching up to 1mm in diameter, and as such we also investigated if optical clearing could be used to achieve increased imaging depths, and so turned to CUBIC optical clearing. As the cultures lack adipocytes, they do not have the same imaging limitations as the *in vivo* gland, and so our clearing protocol was adapted to be limited to a short incubation with CUBIC R2, rather than sequential incubations in R1 which contains high concentrations of detergents. Using confocal microscopy, imaging depths up to 300µm could be achieved, which was considerably greater than that possible with uncleared organoids (typically 100µm). Using this high depth imaging methodology, we noted considerable heterogeneity in organoid morphologies. Organoids could be broadly divided into three morphological types: cystic spheres, branched structures, and budding structures. Interestingly, similar phenotypes have been described in both normal mouse mammary (Jamieson et al., 2017) and human breast cancer organoids (Sachs et al., 2018).

To further profile our organoid cultures, we next assayed expression of normal lineage markers by immunocytochemistry (staining controls for each channel shown in Figure 6.2.1). Basal cells markers in the normal mammary gland include smooth muscle actin protein SMA, the transcription factor p63 and the keratin K5. In the organoids, p63 was seen to be uniformly expressed in the outer cell layers (Figure 6.2.3). SMA was less consistently expressed in both uncleared and cleared organoids, indeed some organoids did not stain well for SMA despite having good p63 expression (Figure 6.2.4a). Given its purpose, SMA could be considered as more

of a functional marker of basal cells while p63 may mark less differentiated cells of the basal lineage. Alternatively, it may be that the process of dissociation and regeneration in an abnormal environment led to SMA expression either being disrupted or not being required, although other studies using the same protocol have found organoids had normal SMA expression (Jardé et al., 2016). Indeed, K5 staining was uniform, although the morphology of some cells was slightly less elongated when compared with that seen in wholemount mammary glands (Figure 6.2.4b). Thus, basal cells are thought to exist in our organoid cultures, despite a lack of SMA expression being seen.

We next examined the expression of luminal lineage markers, E-cad and K8. E-cad expression could be seen uniformly in all luminal cells; however, this was not the case with K8. A number of E-cad+ luminal cells have either very low or no expression of K8. In some organoids, the pattern was particularly interesting – the more budding cultures seemed to have only rare K8+ cells (Figure 6.2.5a). On the other hand, some organoids had K8 expression only in certain regions (Figure 6.2.5b). This is interesting given the variable K8 expression that was seen in the *in vivo* gland (see Chapter 4.2.2.2). The more budded cultures did seem to have a morphology reminiscent of lobuloalveolar structures seen during lactation. Further work examining protein expression would be useful to explore this further, in particular, markers such as β-casein could enable us to investigate the process of lobuloalveologenesis and milk production.

A vital function of the luminal lineage is its response to endogenous hormones and endocrine cues, and as such, expression of hormone receptors is seen as a hallmark of the luminal lineage. Therefore, expression of PR was assessed in the organoids. PR expression was seen in a number of luminal cells (Figure 6.2.6a and b). Together, this suggests that despite aberrant K8 expression in some cells, the luminal lineage is maintained in these cultures. Again this observation reflects that observed *in vivo* where a subset of luminal cells express PR.

To explore stem cell enrichment by our culture conditions, organoids were also cultured in basic media, without the main growth factors that have been shown to be required for mammary gland organoid growth (Nrg1, noggin and Rpso1) (Jardé et al., 2016; Jamieson et al., 2017). Cells cultured under these conditions did grow as small 3D spheres, but at a much slower rate compared to those grown in complete media. After a week, most cultures were mainly composed of stromal cells and fibroblasts. The organoids were examined using confocal microscopy for lineage markers, and although both SMA+ basal cells and K8+ luminal cells were

present, the overall morphology was disrupted, with discontinuous cell layers seen. Moreover, these cultures did not re-grow on passage, suggesting the growth factors used are required for organoid maintenance.

Organoids established from other tissues do not necessarily perfectly reflect the *in vivo* morphology seen in their originating tissues. For example, liver organoids are composed of a single layer epithelium in a cystic sphere, as opposed to a branching structure as seen in the *in vivo* ducts (Huch et al., 2013b). Organoids in our study displayed a branching pattern reminiscent of the *in vivo* mammary gland, although they did not always have a single bi-layer of basal and luminal cells. As such, lumen presence was assessed using the tight junction scaffolding protein, zona occludens 1 (ZO1). In the *in vivo* mammary gland, ZO1 is expressed on the apical membrane of luminal cells, where it is essential for tight junction formation (Stull et al., 2007). In our organoids, ZO1 expression could be seen in multiple areas within the centre of the cultures (Figure 6.2.8). However, there did not appear to be continuous expression around a central lumen, suggesting tight junctions and lumen formation within the organoid is not identical to the *in vivo* mammary gland. Interestingly, in 3D mammary cultures grown in ECM gel, ZO1 expression is continuous on the apical surface of organoids (Jechlinger et al., 2009). This further reinforces the importance of recapitulation of the *in vivo* ECM in these cultures.

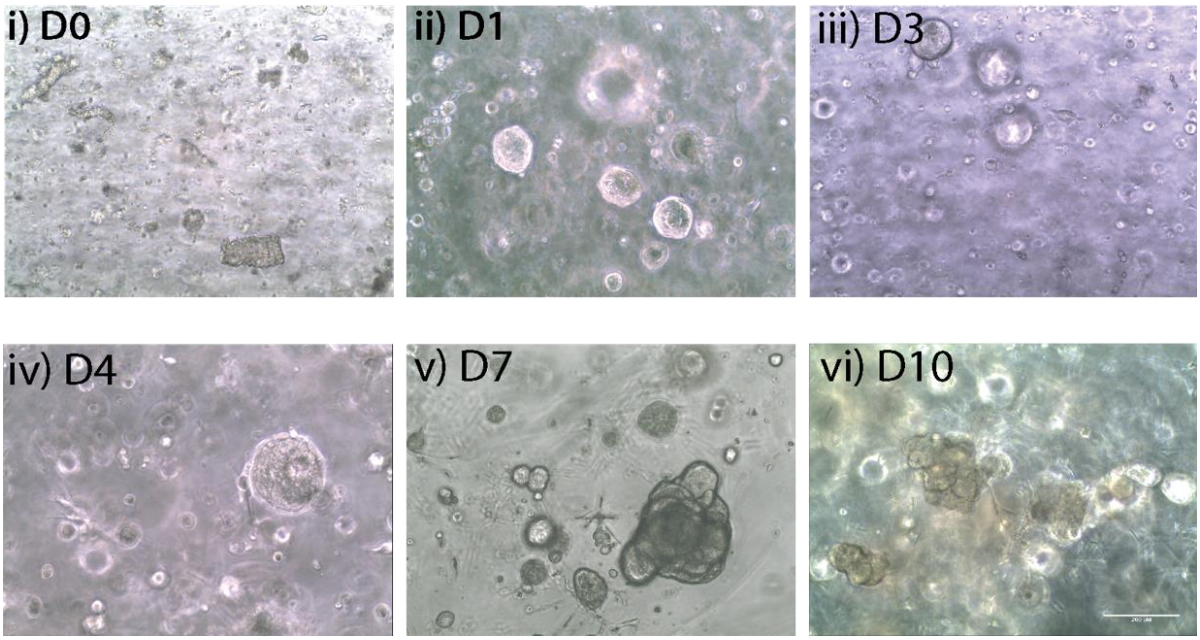


Figure 6.2.1 Time course showing the initial growth of mammary gland organoids

Organoids were cultured in Matrigel and imaged using brightfield microscopy from day of culture (Day 0, **(i)**. By the end of the first day, small spheres had formed which expanded in size **(ii-iv)**. By day 7 **(v)**, organoids began displaying a branched morphology which became gradually more complex, shown up to day 10 **(vi)**.

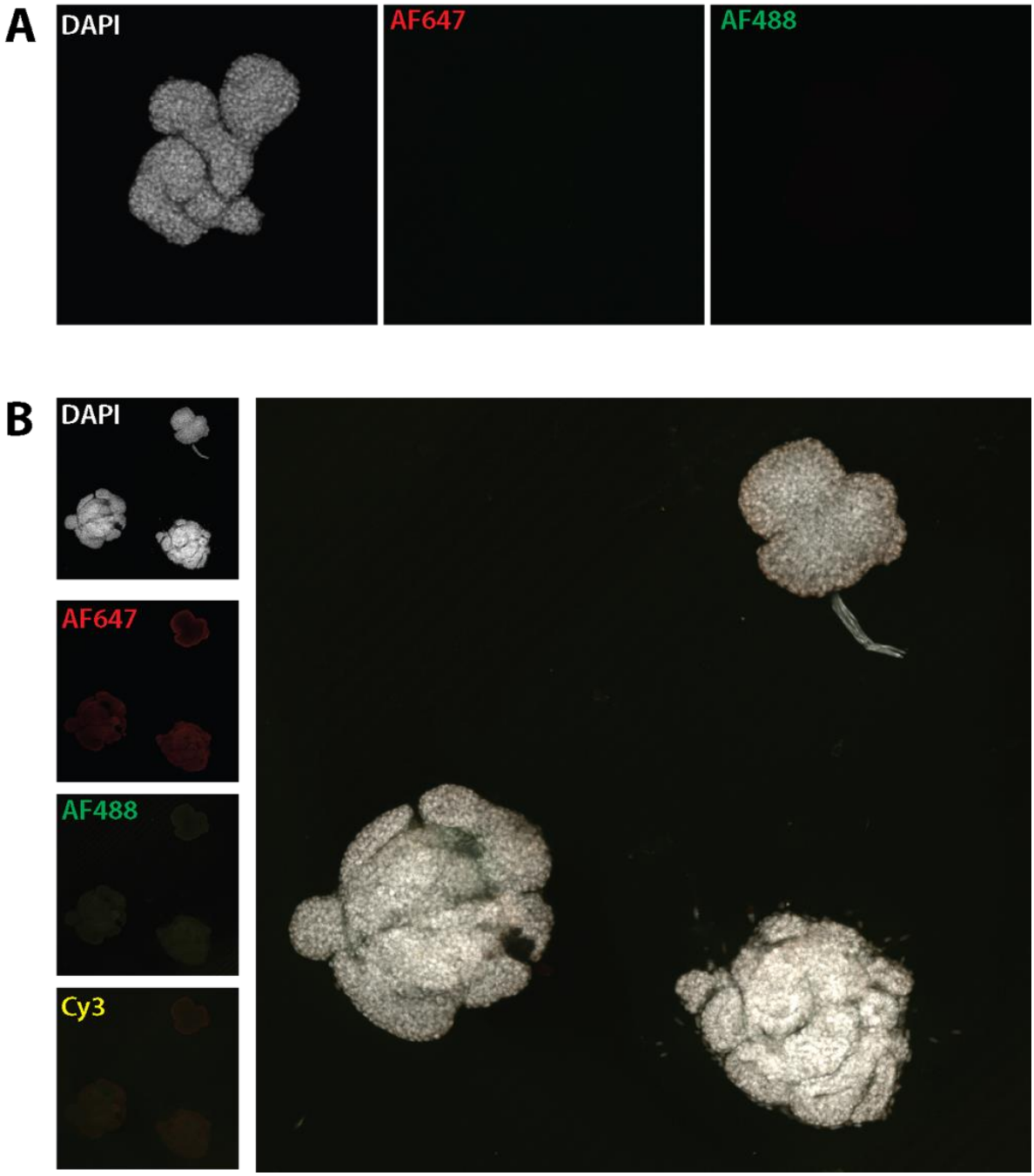


Figure 6.2.2 Staining controls for confocal microscopy of organoids

Organoids were prepared for staining using the normal protocol (see methods), excluding the primary antibody staining step. Images show maximum intensity projections of secondary antibody only controls in: **(A)** wholemount uncleared organoids and **(B)** cleared organoids. Images are maximum-intensity z-projections of individual channels and merged.

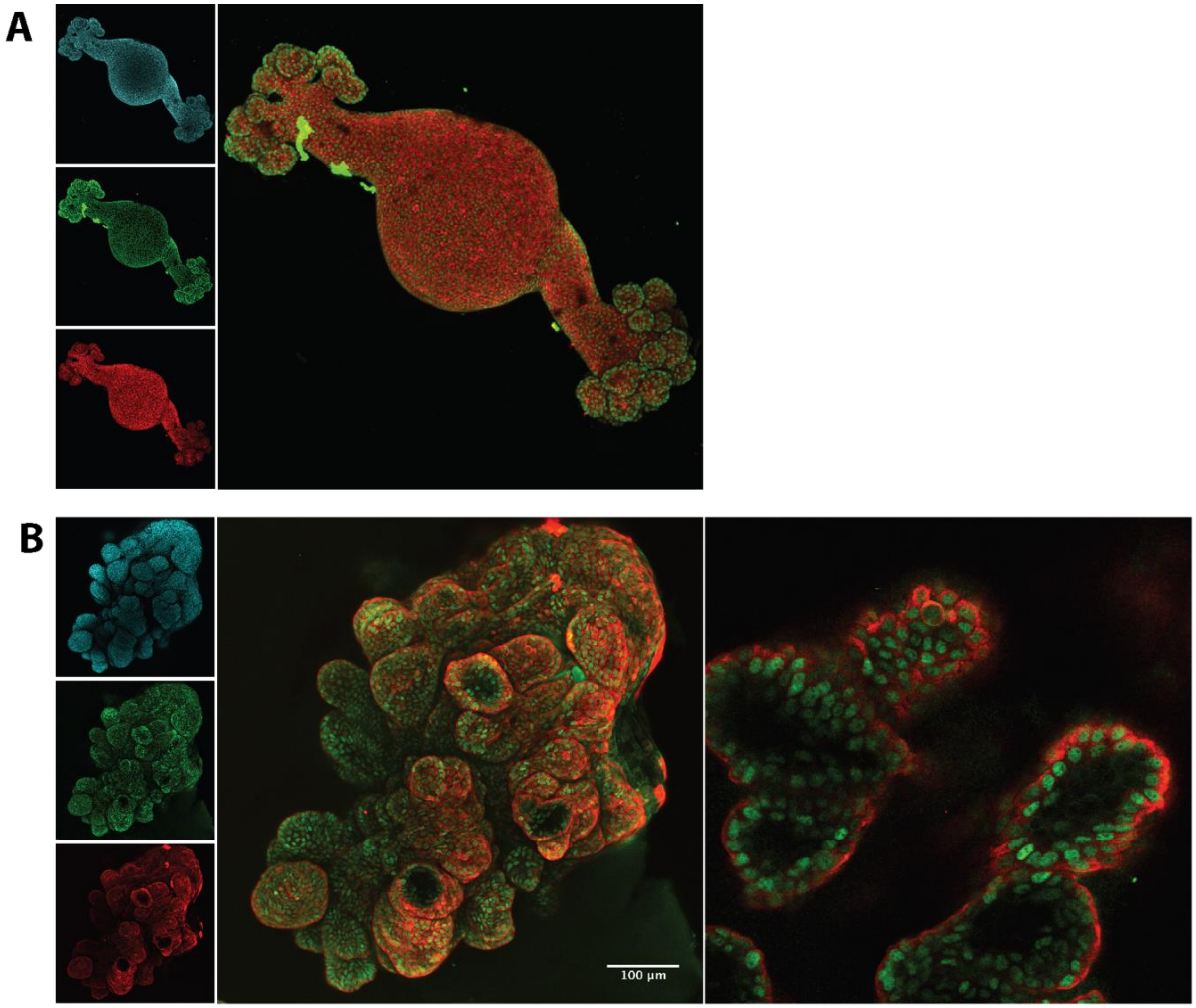


Figure 6.2.3 Basal lineage marker staining in mammary gland organoids: p63

(A) and **(B)** show representative images of p63 expression in basal cells of mammary gland organoids, also stained with DAPI, K8 and E-cad. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

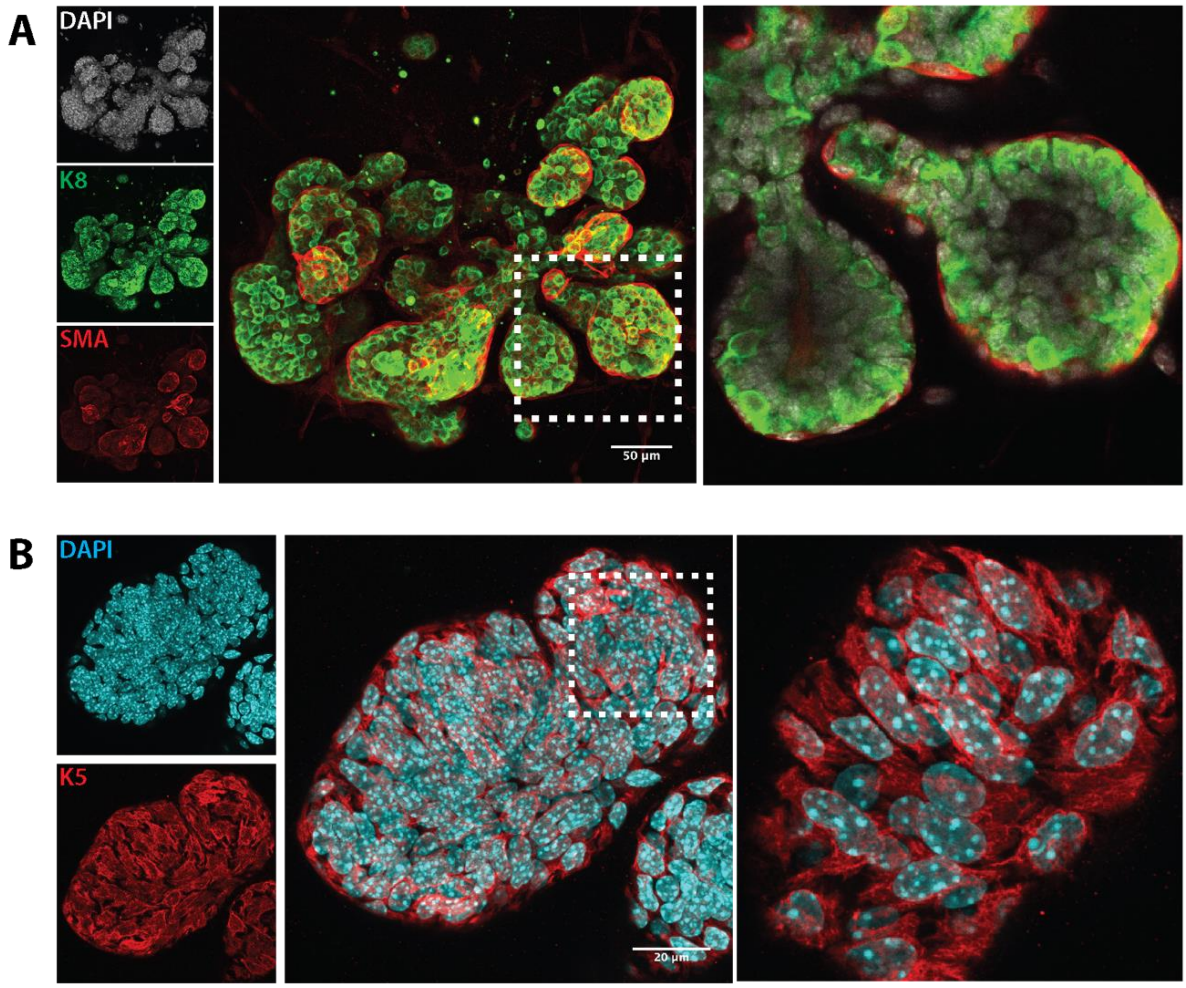


Figure 6.2.4 Basal lineage marker staining in mammary gland organoids: SMA and K5

(A) Representative image showing incomplete SMA expression in basal cells of mammary gland organoids, also stained with DAPI and K8. **(B)** Representative image showing normal K5 expression in basal cells, also stained with DAPI. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged).

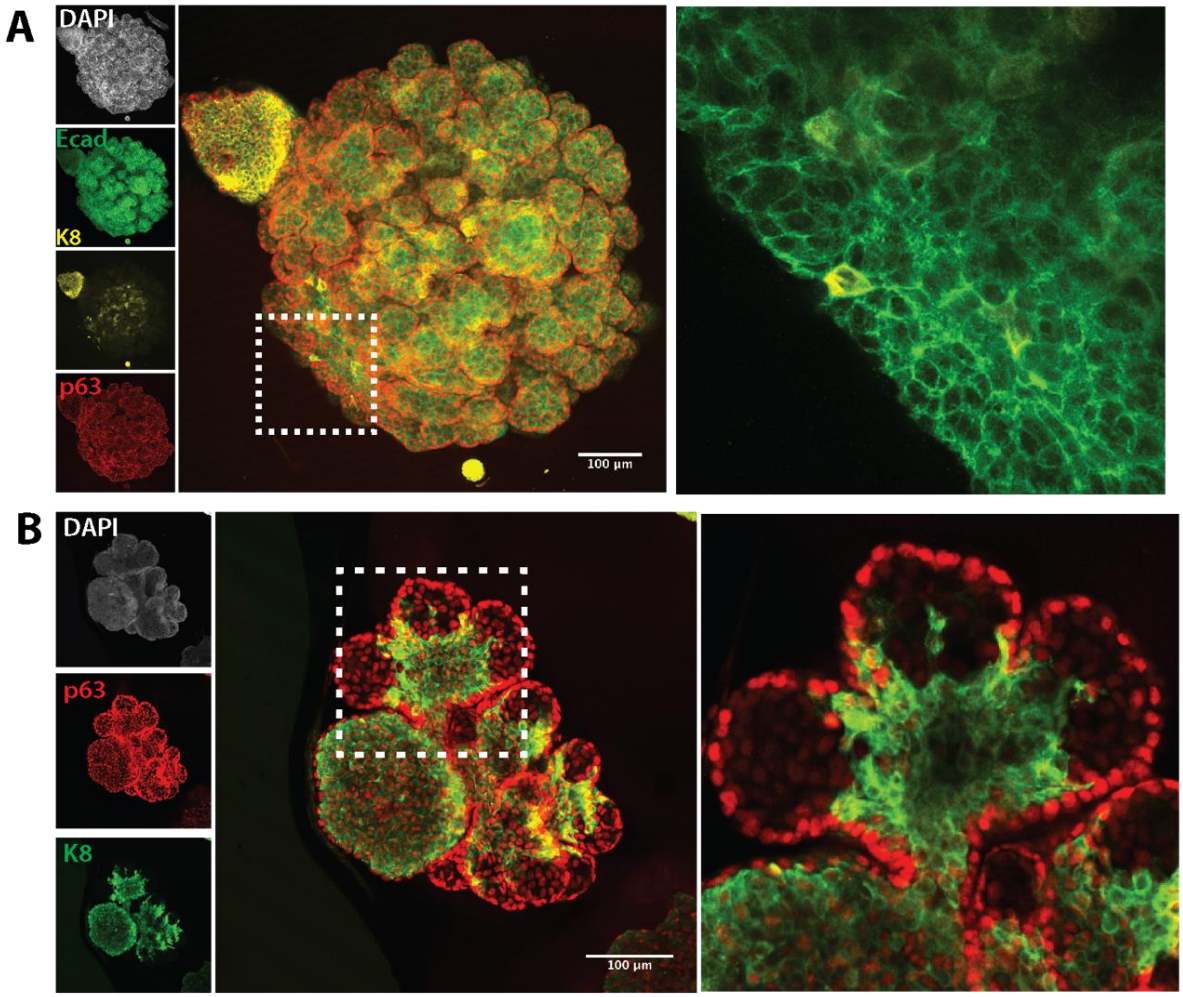


Figure 6.2.5 Luminal lineage marker staining in mammary gland organoids: K8 and E-cad

(A) Representative image showing sparse K8 labelling in luminal cells of mammary gland organoids, also stained with DAPI, K8 and E-cad. **(B)** Representative image of K8 expression in luminal cells only in the centre of an organoid, also stained with DAPI and p63. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

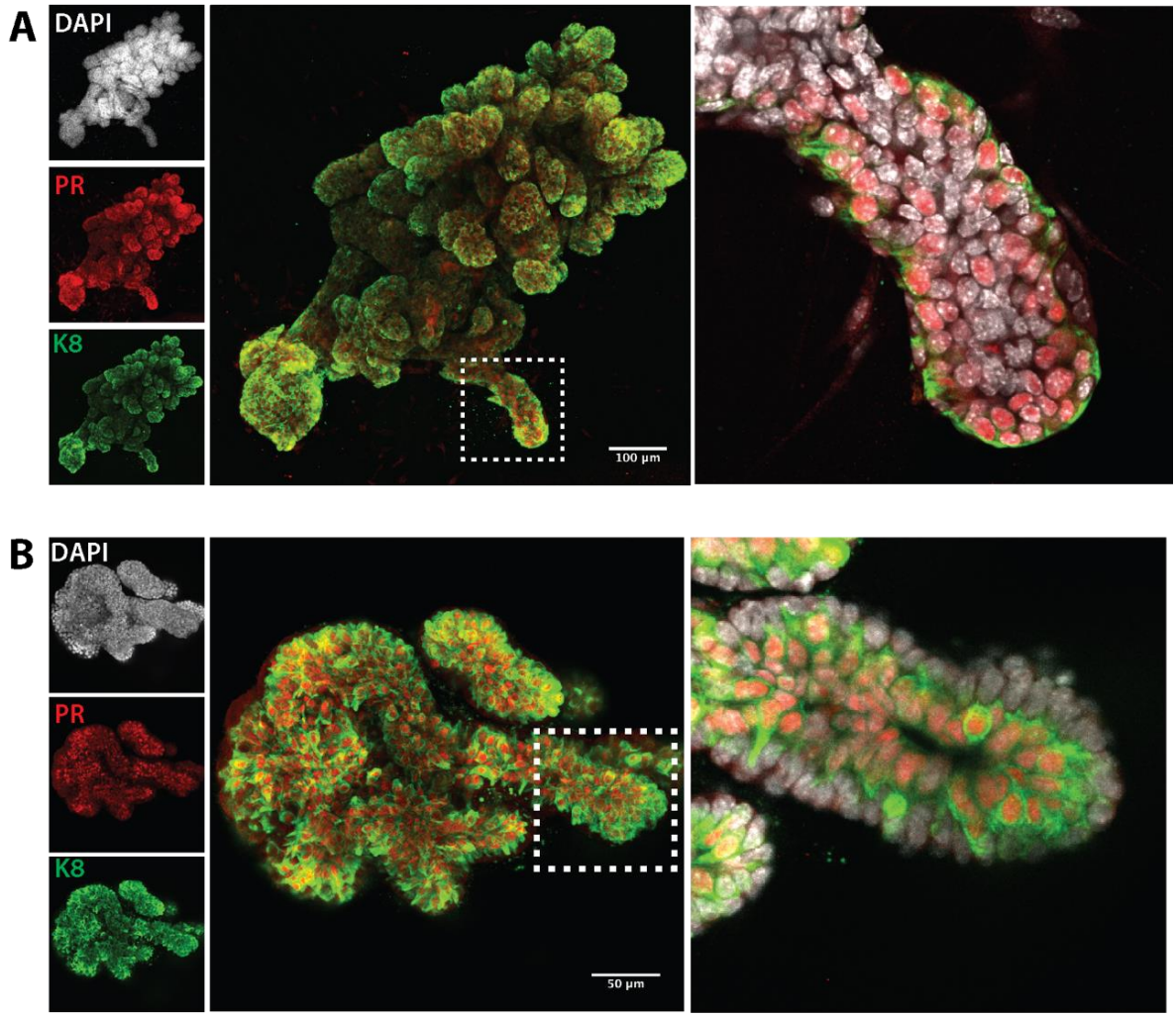


Figure 6.2.6 Luminal lineage marker staining in mammary gland organoids: hormone receptors

Representative images showing PR expression in luminal cells of mammary gland organoids, also stained with DAPI and K8. **(A)** and **(B)** show an inner layer of K8+PR+ luminal cells. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

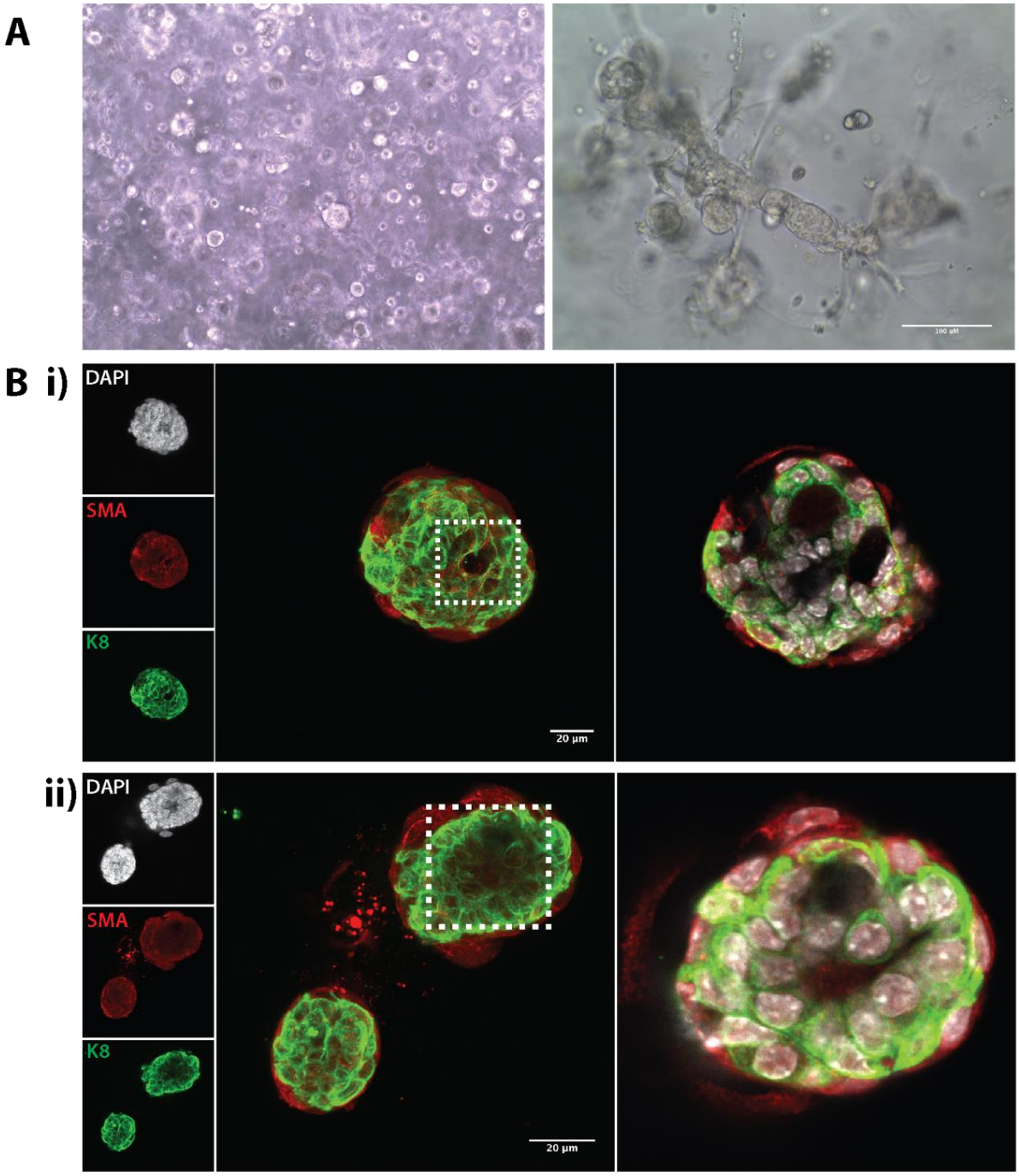


Figure 6.2.7 Mammary gland organoid formation in basal media

(A) Representative brightfield images showing growth of mammary gland organoids after 1 week in culture in basal media, without growth factors. **(Bi,ii)** Representative confocal images showing SMA+ basal and K8+ luminal cells in basal cultures conditions, with discontinuous cell layers seen. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

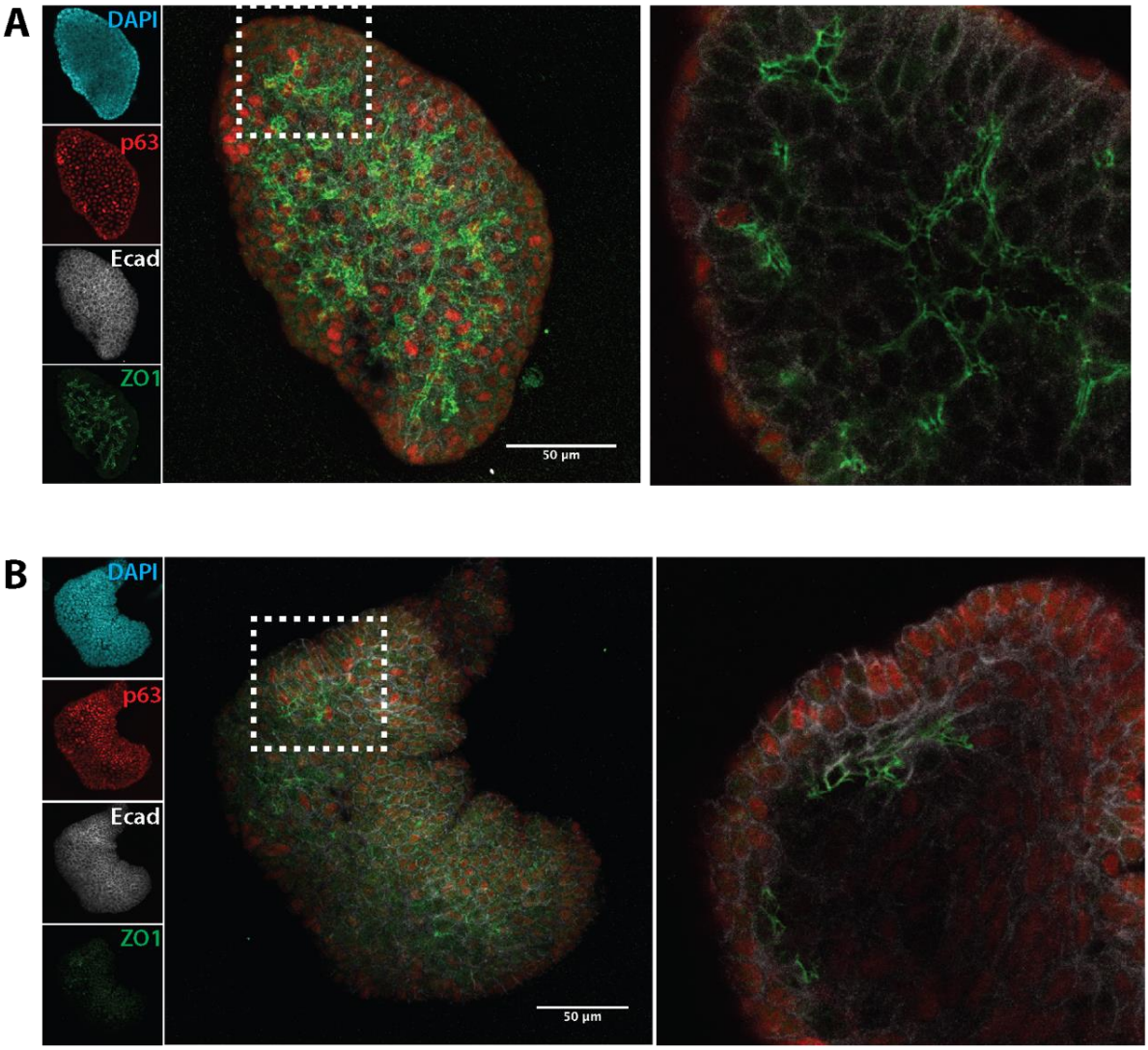


Figure 6.2.8 Lumen formation in mammary gland organoids using ZO1

Representative images showing ZO1 expression in mammary gland organoids, also stained with p63 and E-cad (for basal and luminal cells, respectively). **(A)** and **(B)** show an inner layer of a large organoid. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

6.2.2 Investigating stromal cell presence in organoid cultures

As previously discussed, stromal cells play an important part in mammary gland development and homeostasis (See Chapter 1.1). CD45 staining in optically cleared tissue has shown an abundance of immune cells both in the mammary stroma, and in close proximity to the ductal network (Dr Jessica Hitchcock, unpublished, Figure 6.2.9). CD45 is a pan-immune cell marker, selected to detect the presence of any of the several immune cell subtypes shown to be present in the mammary gland (see introduction). The organoid culture method used in this study was selected in part to maintain any native stromal cells by avoiding specific depletion steps.

Indeed, although the growth factors included in this culture are specific for mammary epithelial growth, brightfield images of the cultures showed a number of what appeared to be stromal cells growing, in addition to the mammary organoids, and often in close proximity (Figure 6.2.10). Immunocytochemistry confirmed the presence of multiple CD45+ cells in our organoid cultures, with a number also seen in tight association with the epithelial cells, in a similar manner to that seen *in vivo* (Figure 6.2.11). Further examination of the morphology of these cells suggested that several different immune cell subtypes may be present, although further staining would be required to accurately identify which immune cell types these are. Indeed, some small, rounded cells with an appearance similar to T-cell morphology (Figure 6.2.12a, arrowhead), whereas other larger cells were reminiscent of monocytes such as macrophages (Figure 6.2.12b, arrowhead). The purpose of these cells in the organoid cultures is not clear – whether they support organoid development or rather are passively remaining from mammary gland dissociation is yet to be seen - functional studies or investigation into the cytokine environment of these cultures would certainly be of value.

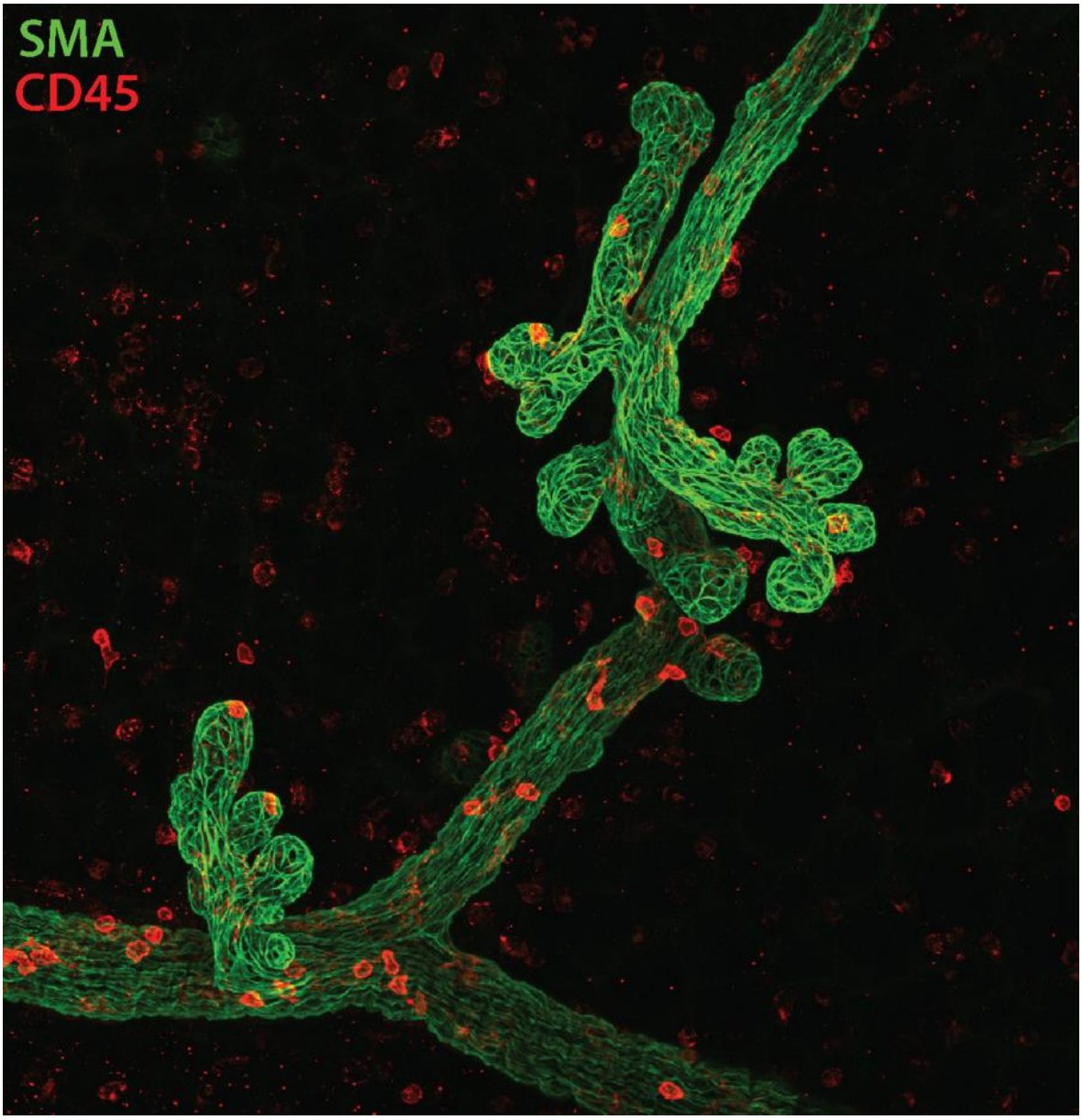


Figure 6.2.9 Example of immune cells in the in vivo mammary gland using CD45

Maximum-intensity z-projection of CD45-expressing immune cells in an adult virgin mammary gland with SMA co-staining, imaged in 3D using CUBIC optical tissue clearing and confocal microscopy. Image courtesy of Dr Jessica Hitchcock (unpublished).

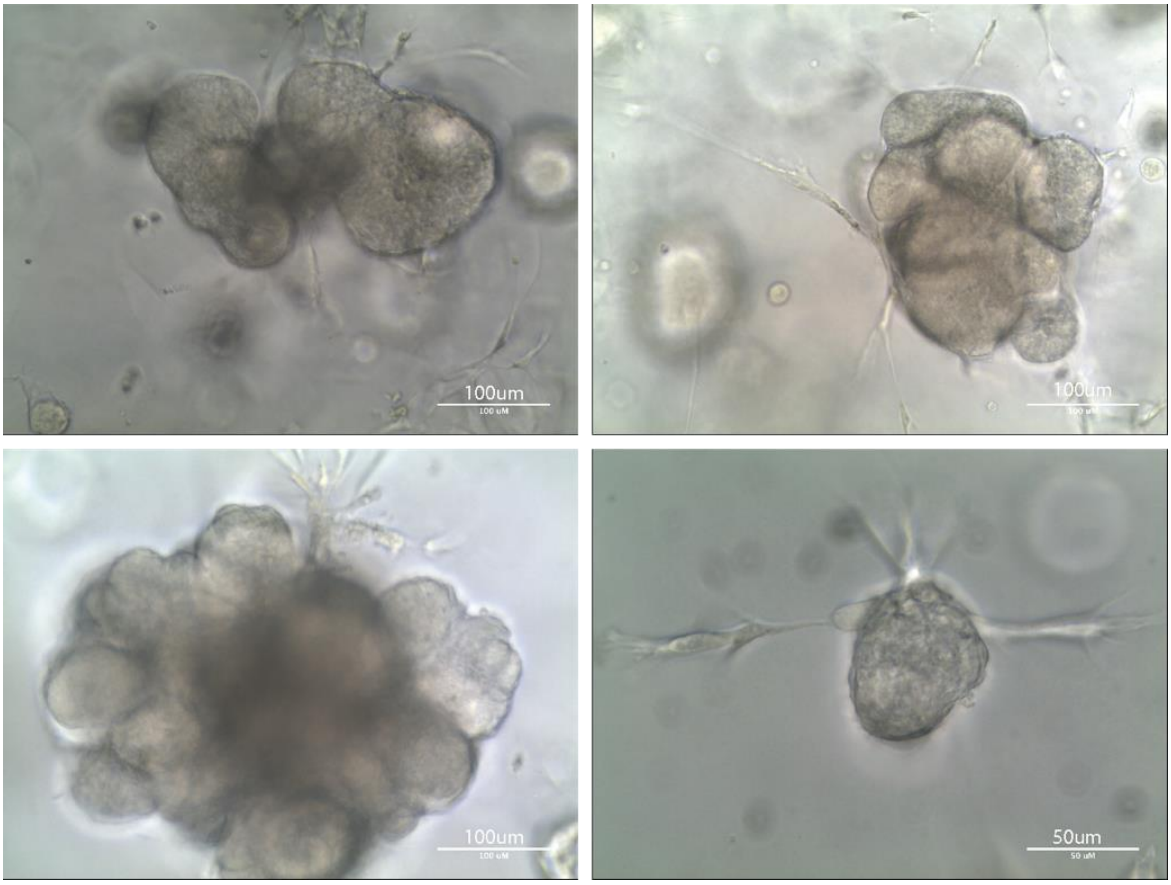


Figure 6.2.10 Brightfield images of stromal cells in mammary organoid cultures

Representative images showing examples of murine mammary gland organoid cultures containing stromal-like cells in close contact with the epithelial structures.

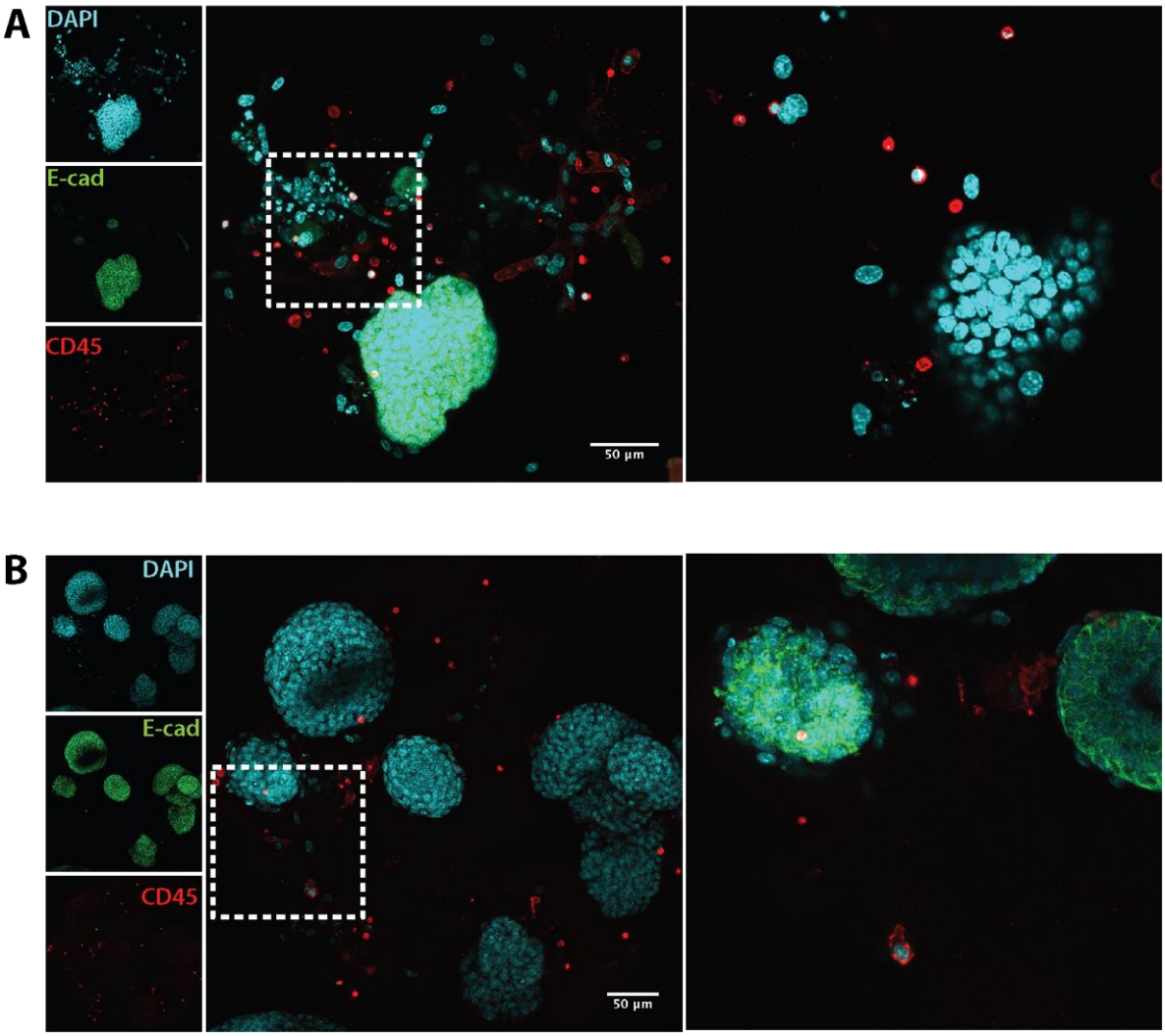


Figure 6.2.11 CD45 cells in mammary gland organoid cultures

Representative images showing examples of CD45+ immune cells in mammary organoid cultures. Cells of different morphologies could be observed with small, T-cell like cells **(A)** seen. Larger cells with a more monocyte-like morphology were also seen **(B)**. Both were in close proximity to the epithelial organoids. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

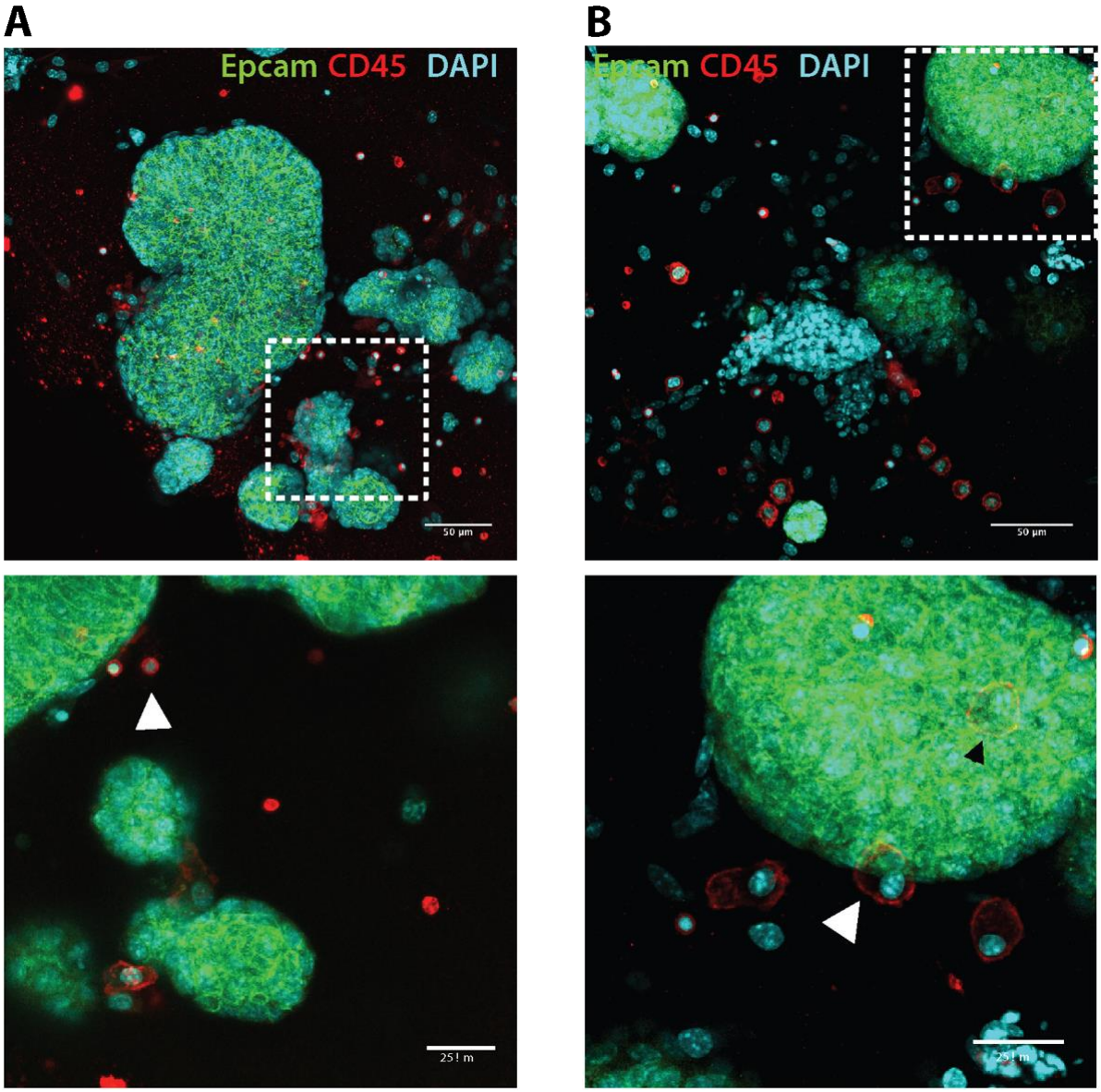


Figure 6.2.12 CD45 cells in mammary gland organoid cultures

Further representative images showing examples of CD45+ immune cells in mammary organoid cultures. Cells of different morphologies could be observed with small, more T-cell like cells **(A)** seen. Larger cells with more of a monocyte-like morphology was also seen **(B)**. Both were in close proximity to the epithelial organoids. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

6.2.3 Investigating proliferation in organoid cultures

In the previous chapter, the proliferation in the mammary gland was investigated using Ki67 staining and EdU incorporation (see Chapter 5.2.2). By examining the mammary gland in multiple focal planes, we uncovered interesting patterns of proliferation not often described in the literature, such as proliferation in the duct during pubertal morphogenesis. As such, we next investigated proliferation in our organoid models to elucidate whether these proliferative patterns were recapitulated *ex vivo*. In particular, we sought to answer whether expansion of the branched organoids can occur through proliferation throughout the organoids - in a similar manner to the *in vivo* gland - or whether expansion occurs due to tip-dependent growth. This was first examined using Ki67 staining, labelling cells in G1/S/G2/M phase of the cell cycle. Ki67+ cells were generally found throughout the organoids, in both basal and luminal epithelial cells (Figure 6.2.13 a,b) suggesting proliferation throughout.

Next EdU incorporation in the organoids was used as a more specific marker of cells undergoing DNA replication. In a previous chapter (Chapter 5.2.1), slight differences were seen in Ki67 staining and EdU incorporation in the mammary gland *in vivo*, potentially due to the differences in cell cycle phases they mark. We selected a pulse-chase of two hours in our organoid cultures - slightly shorter to that used *in vivo* in an attempt to account for differences in EdU diffusion through the ECM as opposed to I.P. injection in mice. Cultures were then stained for EdU and lineage markers, including an EdU staining control (Figure 6.2.14), and imaged using confocal microscopy.

Generally, EdU incorporation was seen to occur in a similar pattern to Ki67 staining, with specific differences seen depending on the type of organoid morphology. In budding organoids, which had shorter protrusions, EdU+ cells were seen throughout the organoid (Figure 6.2.15). In more branched cultures, the majority of EdU+ cells appeared to be found on the tips of the branches (Figure 6.2.16). EdU incorporation occurred in both luminal and basal cells within the organoids (Figure 6.2.17). Ongoing work should focus on quantification of proliferative regions and specific differences between organoid structures in order to further characterise the system.

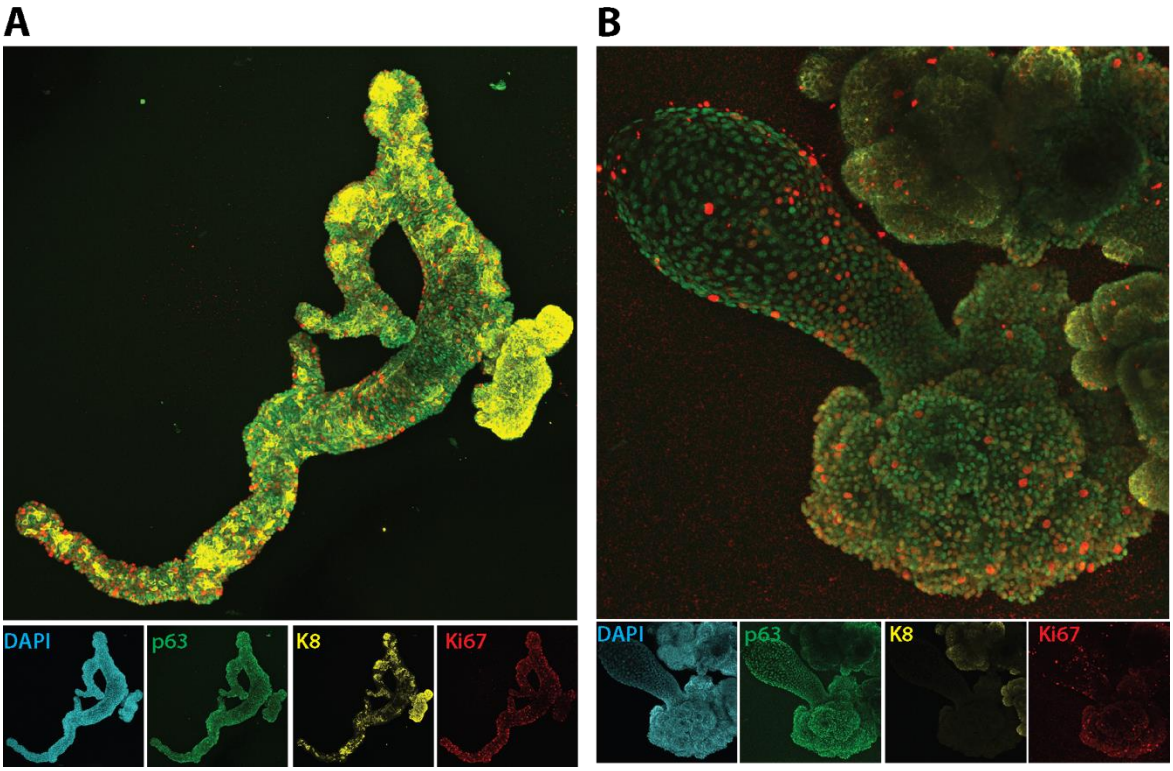


Figure 6.2.13 Examination of proliferation in mammary organoids using Ki67 staining

Representative images showing examples of Ki67 staining showing proliferation in mammary gland organoids co-stained with DAPI, p63 (basal) and K8 (luminal) lineage markers. Images are maximum-intensity z-projections of individual channels and merge.

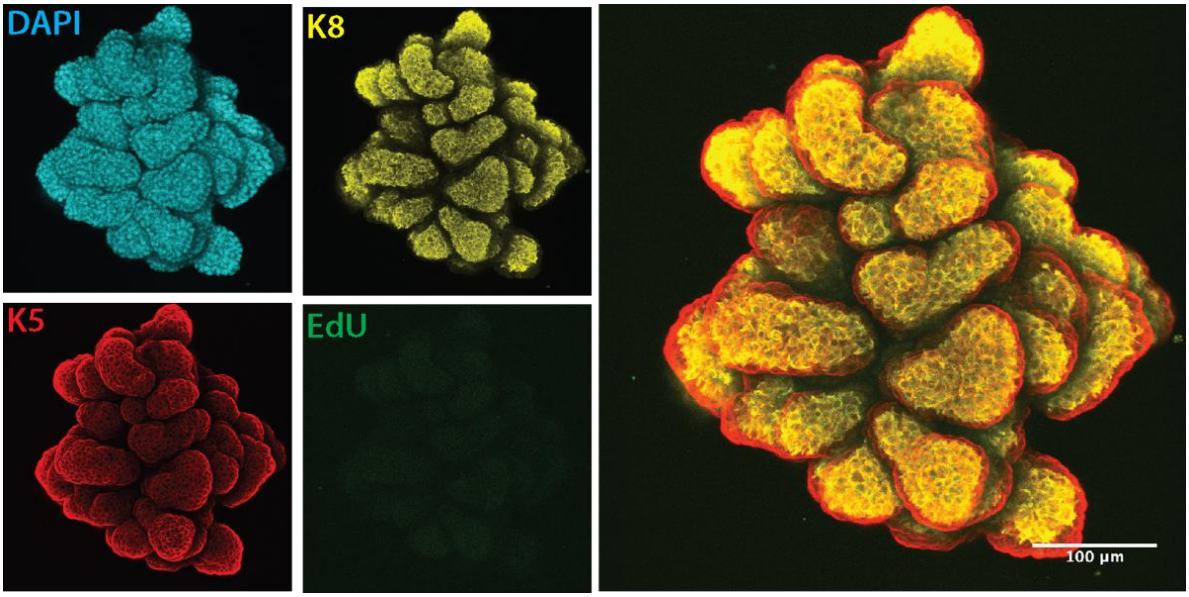


Figure 6.2.14 EdU staining control in mammary organoids

Representative images showing background EdU staining in organoids not treated with EdU, and stained with the EdU staining protocol and DAPI, K8 and K5 lineage markers. Images are maximum-intensity z-projections of individual channels and merged.

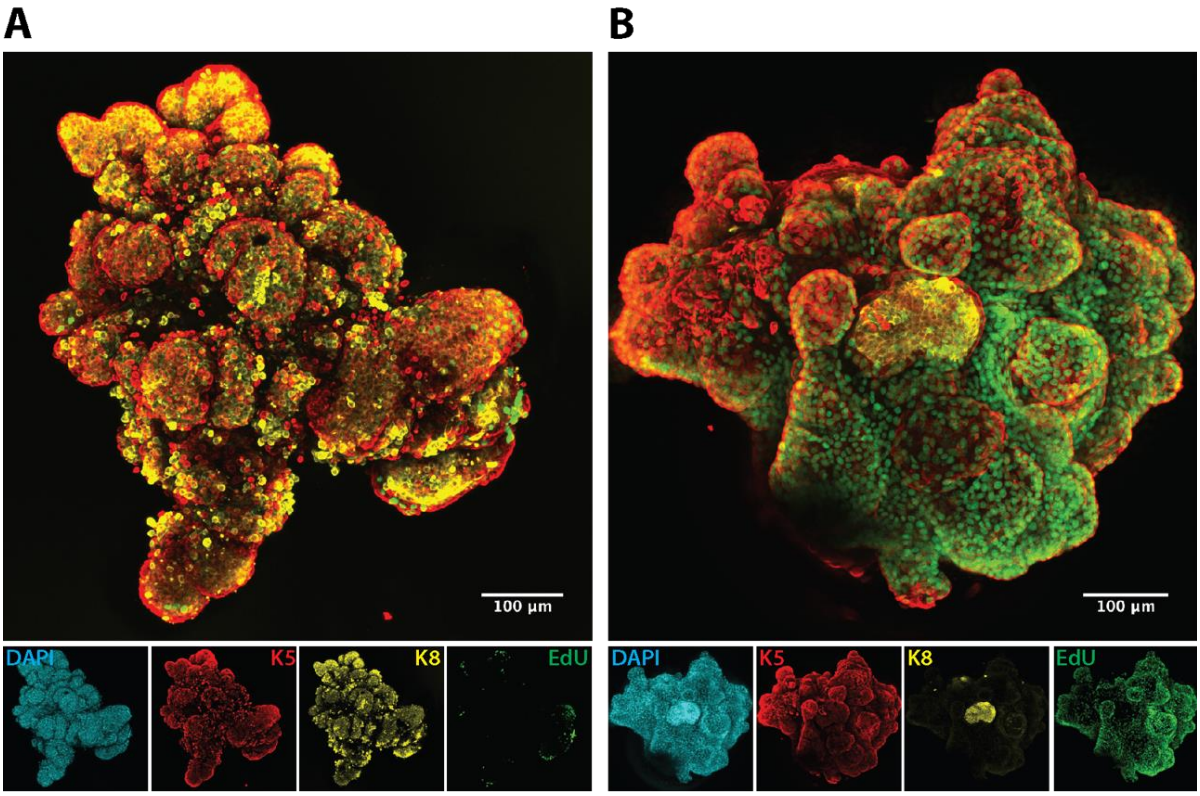


Figure 6.2.15 Examination of proliferation in mammary organoids using EdU incorporation

Representative images showing proliferation in organoids by EdU incorporation, co-stained with DAPI, K8 and K5 lineage markers. Images are maximum-intensity z-projections of individual channels and merged.

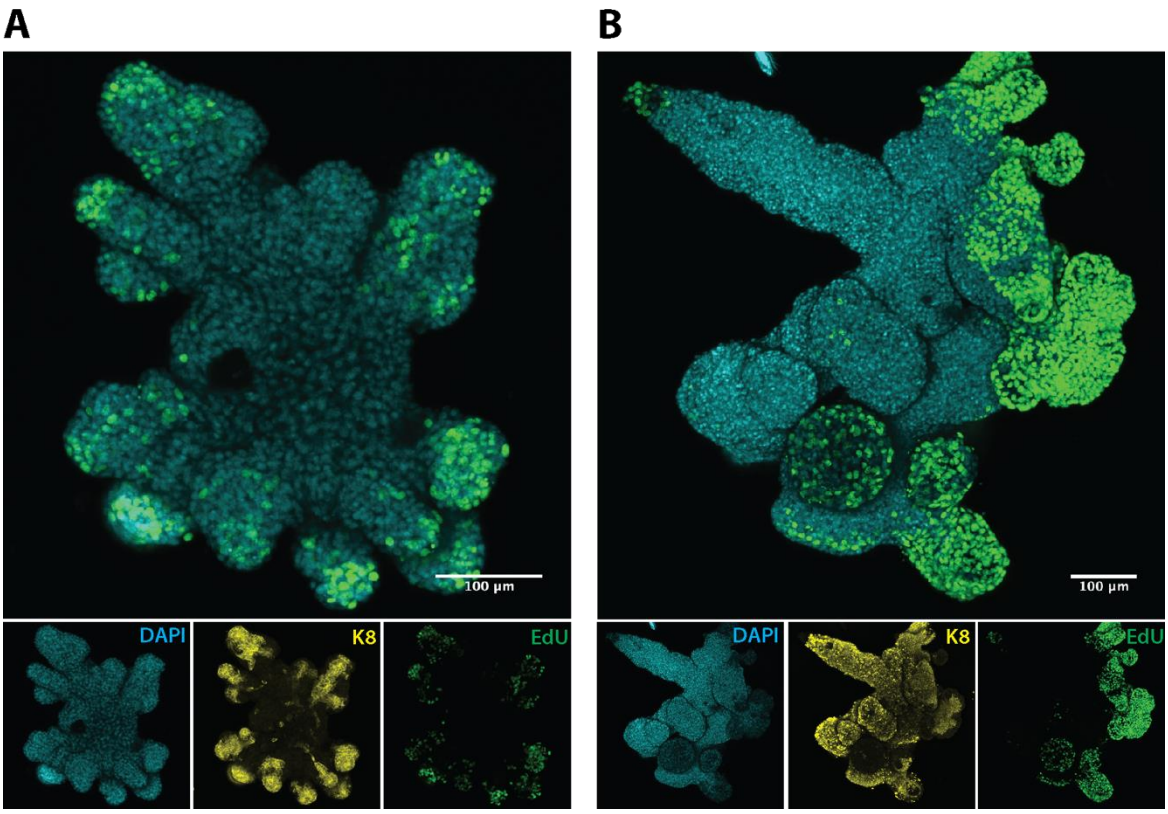


Figure 6.2.16 Examination of proliferation in mammary organoids using EdU incorporation

Representative images showing proliferation in branched tips of organoids by EdU incorporation, co-stained with DAPI and K8 luminal lineage marker. Images are maximum-intensity z-projections of individual channels and merged.

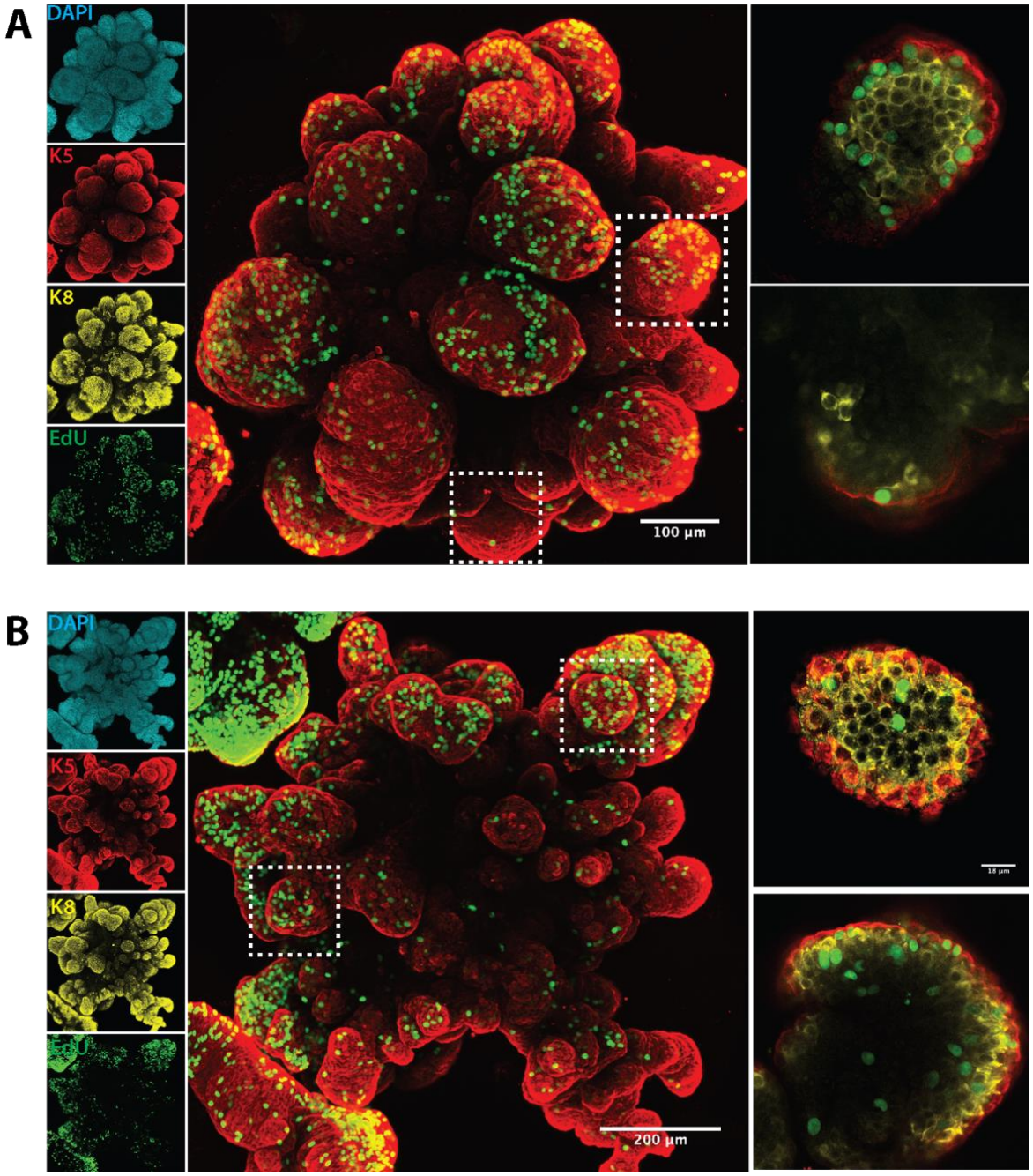


Figure 6.2.17 Examination of proliferation in mammary organoids using EdU incorporation

Representative images showing proliferation in organoids by EdU incorporation, co-stained with DAPI, K8 and K5 lineage markers. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

6.2.4 Lineage tracing

A key area of study in mammary gland biology is defining the specific potency of stem and progenitor cells. Using neutral lineage tracing models, in Chapter 4 we demonstrated that the stem cells involved in the morphogenic events of the mammary gland are lineage restricted. We sought to ascertain whether this conclusion was supported by MaSC growth in our organoid system. To do this, organoids were derived from the mammary gland of mice hemizygous for $R26^{[CA]30}$ YFP and tracing examined at different time points.

Interestingly, despite high amounts of proliferation required for organoid formation, slippage was still relatively infrequent. The [CA]30 slippage rate in murine embryonic fibroblasts in culture (with 95% confidence interval) has been reported as $1.06 \pm 0.38 \times 10^{4}$ per mitosis (Kozar et al., 2013), and although quantification in our organoid system would be needed to confirm this observation, it appears slippage events are rare enough in our model to ensure clonal labelling.

Previously, dissociation of the mammary gland, especially to single cells for FACS, has been shown to significantly disrupt its normal biology – potentially putting it under a "regenerative state" (Shehata et al., 2012). As such, we next investigated if this is the case for this organoid system by examining the potency of clones, characterised by co-expression of lineage markers (SMA, K5 or p63 for basal, and K8, E-cad or EpCam for luminal). All smaller clones examined, comprised of less than 10 cells, were of a single lineage – that is either consisting of only luminal (Figure 6.2.18a,i and ii) or only basal labelled cells (Figure 6.2.18b). Given that [CA]30 slippage in this model is continuous, it may be that these cells represent the progeny of a recently labelled cell, or a progenitor with limited differentiation capacity. This was also the case for medium-sized clonal regions, classed as those with 10 – 50 labelled cells (Figure 6.2.19).

However, a number of larger clones (comprising of more than 50 cells) had labelled cells of both luminal and basal lineages (examples seen in Figure 6.2.20 and Figure 6.2.21). Given both the artificial system, and the regenerative state that may be induced, it is not surprising that mammary stem or progenitor cells involved in organ formation may not be lineage restricted, and in fact be bipotent. Alternatively, it cannot be ruled out that these larger clones represent the progeny from two (or more) lineage restricted stem cells. Whether these cells are the same cells that are involved in ductal elongation and alveologenesis, or whether they are an

entirely different population of stem/progenitor cells that is activated upon culture remains to be determined.

Of note, there was a striking similarity in labelling patterns seen in the organoids to that seen in the *in vivo* mammary gland. Specifically, labelled cells were seen to be intermixed with unlabelled cells. This suggests that these organoids are formed from multiple stem and progenitor cells, in a similar manner to that demonstrated in the *in vivo* gland in Chapter 4. Alternatively, [CA]30 slippage may have occurred in the stem/progenitor cell at a later stage of organoid growth.

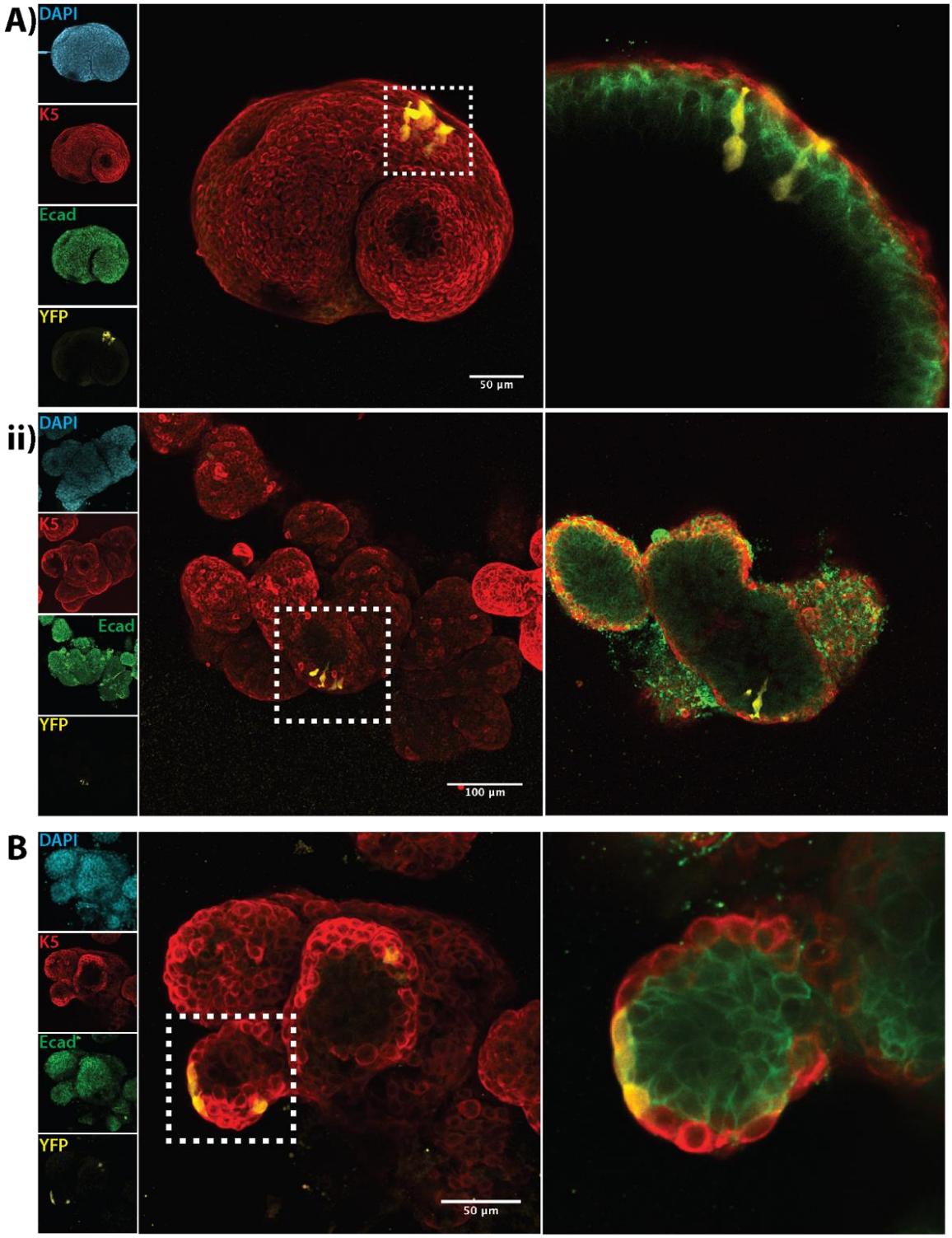


Figure 6.2.18 Examples of small clonal regions in R26$^{[CA]30}$ YFP mammary organoids

Representative organoids showing small clonal labelled regions in R26$^{[CA]30}$ YFP mammary gland organoids, comprised of less than 10 YFP+ cells. (Ai,ii) shows lineage restricted luminal clones. **(B)** shows lineage restricted basal clone. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

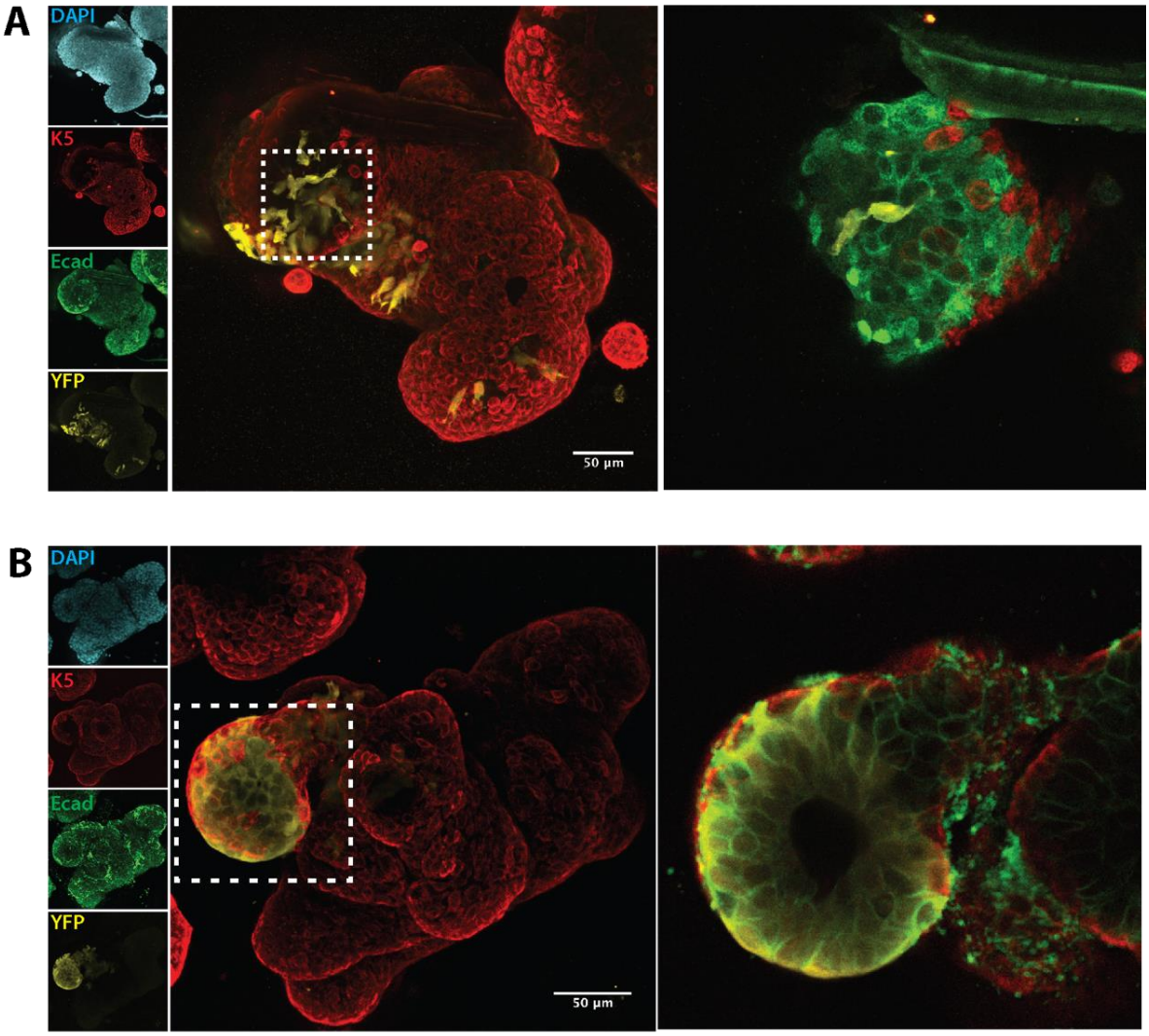


Figure 6.2.19 Examples of medium-sized clonal regions in R26$^{[CA]30}$ YFP mammary organoids

Representative organoids showing medium-sized clonal labelled regions in R26$^{[CA]30}$ YFP mammary gland organoids, comprised of 10-50 YFP+ cells. **(A)** and **(B)** show lineage restricted luminal clones. Images are maximum-intensity z-projections of individual channels, merge and magnified optical slices of a region of interest (boxed and enlarged, right).

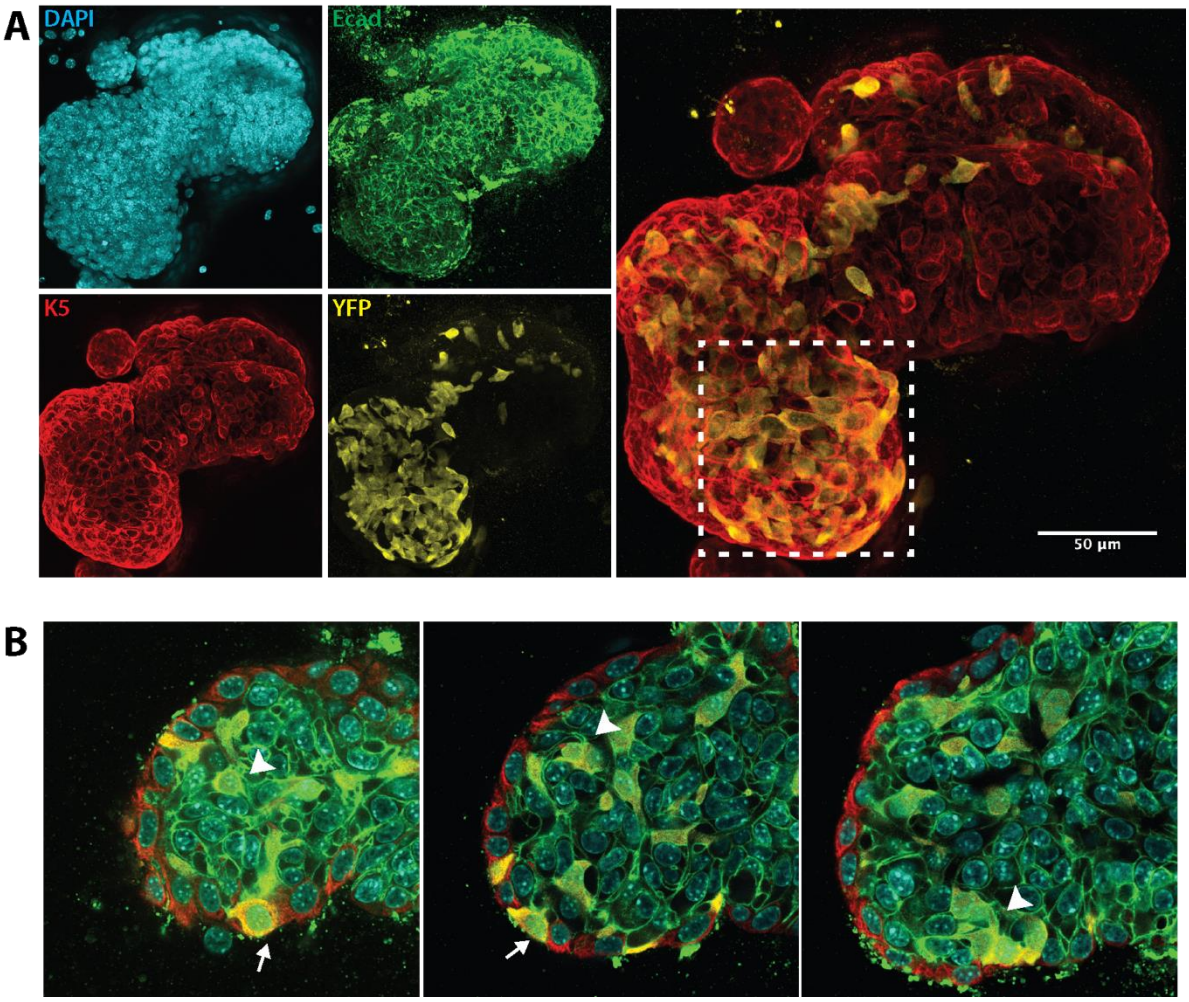


Figure 6.2.20 Example of a large clonal region in R26$^{[CA]30}$ YFP mammary organoids

Representative image of a R26$^{[CA]30}$ YFP mammary gland organoid with a large (50+) labelled clonal region. **(A)** Shows maximum-intensity z-projections of individual channels and merge. **(B)** Magnified optical slices of a region of interest boxed in **(A)**. Labelled cells are both luminal (arrowhead) and basal (arrow).

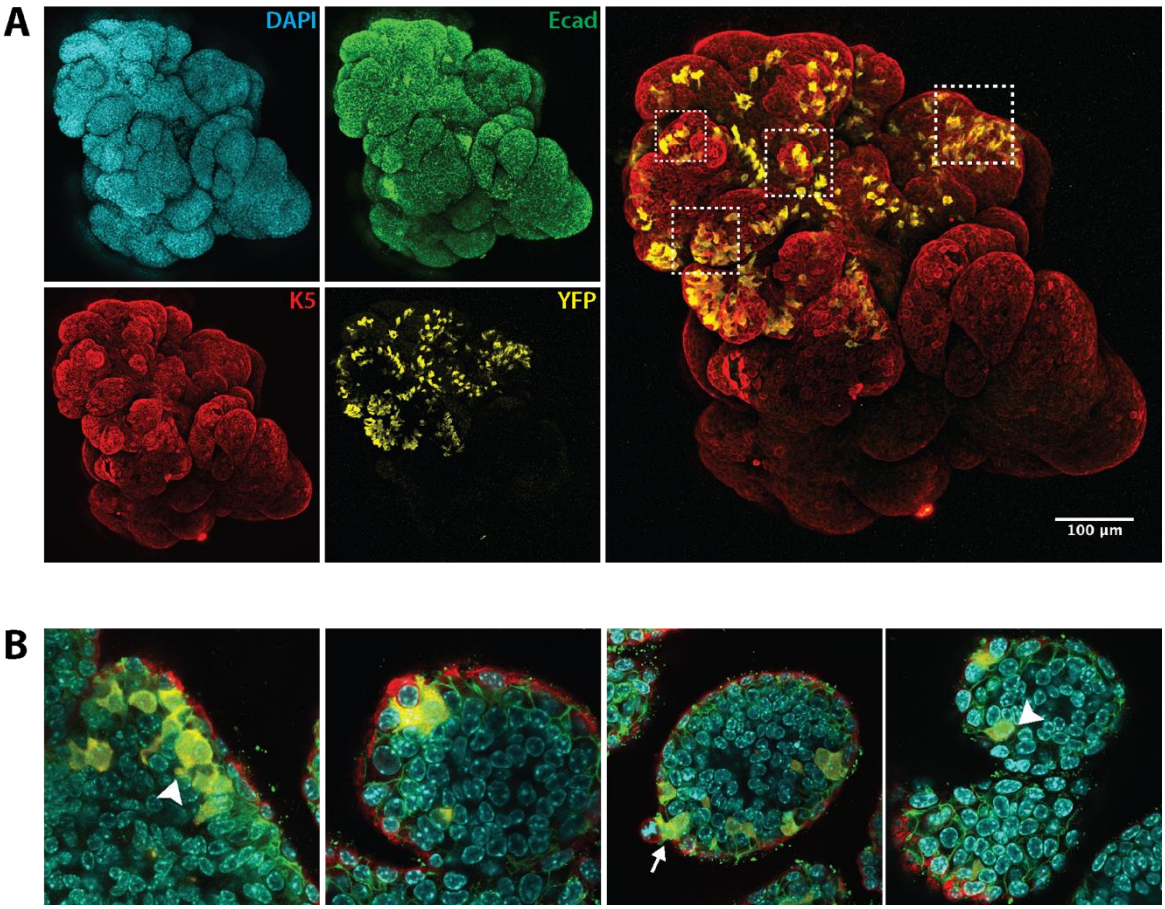


Figure 6.2.21 Example of a large clonal region in R26$^{[CA]30}$ YFP mammary organoids

Representative image of a R26$^{[CA]30}$ YFP mammary gland organoid with a large (50+) labelled clonal region. **(A)** Shows maximum-intensity z-projections of individual channels and merge. **(B)** Magnified optical slices of a region of interest boxed in **(A)**. Labelled cells are both luminal (arrowhead) and basal (arrow).

6.3 Discussion

6.3.1 Overview

Traditional 2D culture systems have multiple drawbacks, specifically in recapitulating the complex milieu of growth factors, cytokines and physical signalling cues found *in vivo*. Although potentially more faithful to the underlying biology, *in vivo* models have specific limitations in their lack of amenability to high throughput screening and ease of perturbation studies. Complex, 3D, culture methodologies are increasingly being developed to bridge the gap between tractability of *in vitro* systems and biological relevance of *in vivo* models. In mammary gland and stem cell biology, the specific utility of such 3D *ex vivo* tissue models are being increasingly understood. However, it is important to understand how well they recapitulate the normal mammary gland biology, together with their specific limitations. The focus of this thesis has been mammary gland development and associated stem cells. In this chapter we aimed to characterise an *ex vivo* culture methodology, specifically focussing on the degree to which our prior findings could be recapitulated in this model.

Our organoid culture protocol was adapted from a methodology first described by Jardé *et al*. In this study, the authors reported that organoids could be maintained in culture for up to two and a half months whist retaining normal morphology and chromosomal stability. In our study, chromosomal number was not specifically investigated, but disruption in the normal cell layers was seen in organoids cultured for up to four months, Interestingly, we noted lack of K8 and SMA expression in some organoids; the exact reason for these differences are not clear, but batch to batch variations can be particularly significant in Matrigel and growth factors. One reason for the absence of SMA may be due to the lack of appropriate basement membrane. Although Matrigel expresses many of the proteins also seen in basement membranes, it cannot fully recapitulate both the molecular and physical characteristics. Given that basal cells, by definition, line the basement membrane, it may be that they are especially sensitive to changes in this.

Matrigel was utilised for culturing of organoids. However, a number of limitations of Matrigel have been highlighted in recent years, largely due to its origin. Matrigel is derived from the secretions from tumours of the Engelbreth-Holm-Swarm (EHS) mouse sarcoma cell line. This means the synthesis is not tightly controlled resulting in significant batch-to-batch variations in exact constitution and protein and growth factor content. Here, the growth factor reduced variant has been utilised, however even this has been shown to vary up to 50% in protein content

among different suppliers (Hughes et al., 2010). This uncertainty of exact composition may result in experimental variations which can lead to reduced reproducibly. Consequently, using Matrigel-based organoid systems such as the one described herein may not be appropriate in experiments which require knowledge of exact protein constituents and concentrations. Here, we have tried to control for these aspects by ensuring experiments used similar batches. However, development of models utilising known ECM substitutes would be valuable.

We sought to determine whether our organoid models recapitulated observations in specific patterns of proliferation seen *in vivo*. Ki67 staining and EdU incorporation studies demonstrated that proliferation largely occurs in the expanding tips of branches within organoids, in a manner reminiscent of TEBs in the pubertal gland (Paine and Lewis, 2017). Indeed, organoid branches could be seen to have bifurcated, highlighting the similarities to the *in vivo* gland. However, Ki67 staining and EdU incorporation only allow investigation of proliferation at a fixed timepoint and in fixed samples. An interesting future study could derive organoids from transgenic FUCCI mice, as shown in Chapter 5, to further investigate proliferation in organoid models. Similar studies have been completed in organoids derived from both the taste buds (Aihara et al., 2015) and intestine (Carroll et al., 2018), which allowed more accurate resolution of cell cycle stage. Further, FUCCI-organoids could be imaged live for real-time assessment of cell cycle stages, giving us an even greater understanding of specific kinetics of proliferation within the growing organoids.

Live imaging could also be utilised to examine lineage tracing *in situ* within the organoids as it happens, which would also allow for definitive examination of potency. Indeed, this was attempted with the $R26^{[CA]30}$ YFP system to examine live slippage and reporter expression. However, unfortunately the fluorescent intensity of the YFP reporter of this model is not strong enough, and imaging required an antibody to amplify the signal. An interesting future investigation could include the use of a reporter with strong enough signal to allow for live imaging. One such system could be the R26-Confetti system, used in Chapter 4. Indeed, the use of this has been demonstrated in organoid systems of both the mammary gland (Jamieson et al., 2017) and taste bud (Aihara et al., 2015). It also has the added benefit of allowing timing of tracing, and multiple colours for better separation of clones.

As discussed, the one of the greatest limitations of organoid systems is that they generally lack a functional microenvironment and the associated signalling that this provides. Attempts to reduce

these limitations, for example by using more biologically relevant ECM-matrices with control of physical signalling cues, have shown promising results (Sokol et al., 2016; Miller et al., 2018). However, one of the aspects that is more challenging to recapitulate is soluble cues, which are particularly important for mammary gland biology - such as ovarian hormone signalling, which is a key driver in the morphogenesis of the mammary gland both during development and homeostasis (Ruan et al., 2005; Rajaram et al., 2015).

Here, PR expression was maintained in the luminal cells suggesting hormone sensing is still possible. Previously, hormones associated with lactation, such as prolactin, have been shown to induce milk protein expression, again highlighting that organoids maintain hormone sensing capabilities (Jardé et al., 2016; Jamieson et al., 2017). Since oestrous cycling is another important aspect of mammary gland biology, it would be interesting to investigate if exogenous application of progesterone and oestrogen, to mimic the different endocrine environments of the different oestrous stages, would affect mammary organoid morphology in a similar manner to that seen *in vivo*.

6.3.2 Importance and implications

Development of better *in vitro* systems for tissue culture is vital for not only increasing understanding of normal biology, but also for providing a system for high throughput screening in drug development – an area which has been severely affected by the lack of appropriate models (Cassidy et al., 2016). Moreover, the use of 3D models is a particularly attractive aspect for studying human breast development, given the difficulties and associated differences between woman and mouse. A number of these models have been investigated using different substrates, including a particularly elegant system using floating collagen gels which facilitate the development of structures resembling human terminal ductal-lobular units (Linnemann et al., 2015). Further characterisation of the role of stem and progenitor cells in the development of these models will provide a vital tool to study both normal and malignant human breast biology.

Chapter 7

Discussion

7.1 Context

Aside from gaining a greater understanding of normal tissue development, study of the stem and progenitor cells within an epithelial tissue structure can offer great insight into the processes of disease, notably including neoplastic transformation. The link between stem cell biology and cancer has been demonstrated in epithelial systems such as the intestine (Kozar et al., 2013) and stomach (Barker et al., 2010), in addition to haematopoietic stem cell involvement in blood malignancies (Prick et al., 2014). Within the mammary gland, the stem cell hierarchy is less well understood. Consequently, the involvement of MaSCs in malignant transformation is not entirely clear, but a number of studies have demonstrated a potential role (Koren et al., 2015; Lawson et al., 2015; Van Keymeulen et al., 2015).

Indeed, the intrinsic subtypes of breast cancer (Luminal A, Luminal B, HER2-enriched and Basal-like) have each been linked to neoplastic transformation of different cells in the MaSC developmental hierarchy, with relatively low concordance amongst the literature (Sorlie et al., 2003; Prat and Perou, 2011). For example, it has been variably suggested that the cell of origin of both sporadic and BRCA1-mutated Basal-like breast cancers is a luminal progenitor cell, but also that an unknown MaSC (not a luminal progenitor) is the cell of origin of all Basal-like tumours (Prat and Perou, 2011). Clearly, further study is required to elucidate the MaSC hierarchy and how malignant transformation at different developmental stages can lead to breast cancer of distinct clinicopathological subtypes.

Breast cancer is the most common cancer, with an estimated 2.4 million new diagnoses in 2015; indeed, one in 8 women and one in 603 men will develop breast cancer between birth and age 79 years. In England alone, breast cancer accounted for 44,000 new cancer registrations in 2016 and was the most commonly diagnosed cancer (Wang et al., 2016). These observations reinforce the global burden due to breast cancer, and the need for improved outcomes through the development of new clinical strategies which will become increasingly important over the next several decades. This can be achieved through a better understanding of breast cancer biology coupled with advances in prevention, detection and treatment. This includes characterisation of the origins and development of breast cancer, which first requires an understanding of the normal biology of the mammary gland (Wansbury et al., 2011).

Using a low-density, neutral, genetic labelling strategy, combined with a method of imaging that is free of proteolytic digestion, we have provided corroborating evidence of the lineage restriction of proliferative stem/progenitor cells at the three major stages of mammary development: in the late embryo, during puberty and in reproduction. These findings also confirm that remarkable heterogeneity exists within the adult mammary stem and progenitor cell compartment and suggests similar multiplicity within their embryonic precursors. However, whether these adult stem cells work cooperatively or competitively to achieve the required developmental and morphogenetic outcomes in the mammary gland is a particularly interesting aspect that remains unclear.

Importantly, this work has also revealed the remarkable capacity of a single embryonic MaSC to contribute to ductal development, providing unprecedented insights that could only be disclosed by this single-cell approach. It is increasingly hypothesised that certain cancers may arise from reactivation of embryonic developmental programmes in postnatal tissues (Howard and Veltmaat, 2013; Wahl and Spike, 2017). Thus, an elucidation of the full spectrum of stem/progenitor cell populations in both the pre- and postnatal mammary gland is paramount for defining the cellular origin of heterogeneous breast tumours (Sreekumar et al., 2015).

The origin of luminal and basal cell lineages in the mammary gland has been the subject of intense investigation and debate. Work shown in Chapter 4, in addition to recent saturation lineage tracing, single-cell lineage tracing and promoter driven lineage tracing studies have provided support for lineage restriction of MaSCs from late embryogenesis (Lilja et al., 2018; Lloyd-Lewis et al., 2018; Wuidart et al., 2018) into adulthood (Van Keymeulen et al., 2011; Davis et al., 2016; Wuidart et al., 2016; Scheele et al., 2017). The work shown here has also demonstrated the capacity of these unipotent stem cells to contribute to the development of the adult mammary epithelium, while at the same time revealing significant redundancy in the construction of each major duct in puberty (Figure 7.1a) and lobuloalveolar structure in lactation (Figure 7.1b) (Davis et al., 2016; Wuidart et al., 2016; Scheele et al., 2017). However, studies attributing postnatal mammary gland development to bipotent stem cells are still being published. Consequently, this raises the question - what level of evidence will be required to achieve a unifying definition of physiological MaSC potency?

Finally, it is important to emphasise that the lack of evidence for the presence of multilineage clones does not unequivocally show that bi/multipotent stem cells do not exist (Visvader and

Stingl, 2014; Rios et al., 2016). Indeed, although the small number of cells initially labelled in the models shown here permits the indisputable analysis of clonal progeny, it also limits the power to detect and characterise the full spectrum of stem and progenitor cells that are present in the mammary epithelium. For example, quiescent bi/multipotent MaSCs, if they exist, would not be detected by this approach. Thus, it is imperative that the epithelial hierarchy in the mammary gland is rigorously assessed at various developmental stages, using a range of methods, models and systems of analyses.

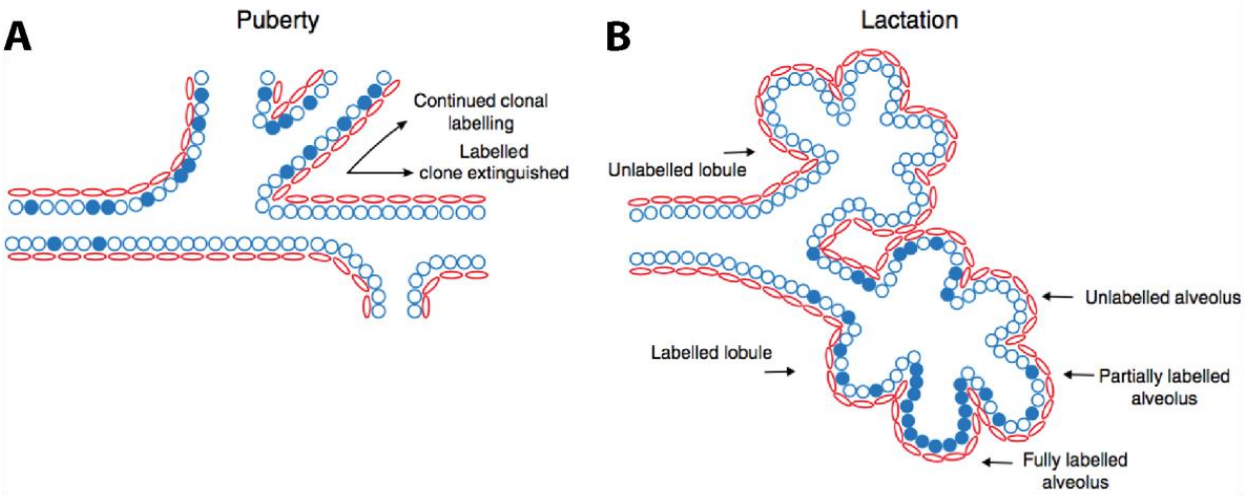


Figure 7.1 Schematic for proposed model of mammary stem cells

Diagram outlines the common labelling pattern arising from the genetic labelling of a single lineage-restricted MaSC/progenitor to **(A)** ductal morphogenesis and **(B)** alveologenesis, identified in *R26*$^{[CA]30}$ mice and confirmed using the R26-Confetti model

7.2 Defining a transcriptional profile of MaSCs

In order to better understand the role of MaSC(s) in health and disease, we must define a possible molecular portrait of the cell. This has been achieved in other adult stem cell systems using single-cell RNA-Seq, which allows whole-transcriptome profiling of individual cells (Bach et al., 2017; Kumar et al., 2017). This can then be combined with methods to either label single cells and all of their progeny (for example from single-cell lineage-tracing) or incrementally mark cells (for example using CRISPR-generated barcoding). This then creates a molecular profile of all cells within a single clone, which may then reveal a transcription profile of the cell of origin.

Significant work has already been done in this regard, but there is still much disagreement in the field and disparities in findings between groups (Lim et al., 2010; Spike et al., 2012; dos Santos et

al., 2013; Bach et al., 2017). Moreover, a definitive marker of MaSCs has yet to be identified by these studies. Indeed, one recent study isolated cells from developing TEBs and found significant transcriptional heterogeneity in both luminal and basal cell populations. This highlights that it may not be possible to link behaviour and identity of stem cells in the mammary gland to a single molecular profile or specific markers (Scheele et al., 2017). Whether this heterogeneity is a function of varied positional location of MaSCs, or a reflection of the multiple states of potential MaSCs is as yet unclear.

Stem cells are defined by their functional abilities, that is: proliferation, self-renewal, the production of a large number of differentiated progeny, tissue regeneration/repair, and flexibility within these states (Potten and Loeffler, 1990). The challenge in the mammary gland thus far has been how to study the functionality of a cell without inadvertently altering its function. Further refinement of lineage-tracing approaches and the application of other novel experimental models and methods for marking, visualising, and profiling individual cells will continue to provide important insights in this field (Clevers and Watt, 2018).

7.2.1 Tracing of embryonic mammary stem cells

A number of studies have previously revealed the bipotent nature of embryonic MaSCs, using both transplantation and lineage tracing studies (Van Keymeulen et al., 2011; Boras-Granic et al., 2014). Here, lineage restriction in embryonic MaSCs from E17.5 was demonstrated using neutral fate mapping models from embryonic mammary development, to pubertal morphogenesis of the gland. This is corroborated by other reports using Notch lineage tracing models (Lilja et al., 2018; Wuidart et al., 2018). The signals involved in this switch from multipotency to unipotency is unclear and questions remain regarding the mechanisms, whether they be intrinsic – through transcriptional control - or extrinsic – for example via microenvironmental signalling. Exploration to further define the normal biology of the gland combined with delineating the exact time point and a molecular understanding for this temporal switch will be useful.

7.2.2 Examination of proliferation.

Using wholemount imaging, we have also shown where the proliferative regions lie in the mammary gland, and that proliferation levels dramatically reduce with age. The finding that EdU incorporation could be detected in the duct was unexpected, having not been previously described. Proliferation and growth in the developing mammary gland is usually reported to be

mediated by the TEBs (Kenney et al., 2001; Hinck and Silberstein, 2005). This supports the hypothesis that MaSCs may be laid down along the duct from the TEBs as the gland develops. Moreover, traditionally it has been suggested that MaSCs reside in the basal population (Plaks et al., 2013; Rios et al., 2014; Wang et al., 2015). Here, we have shown that at least during puberty, proliferation is found in both the luminal and basal lineages, in addition to the cap and body cells in the TEBs.

In the post pubertal gland, oestrous cycling in the mammary gland provides a fascinating model of postnatal epithelial proliferation, degradation and regeneration on a short time scale. Here, proliferation was examined during dioestrous using 3D imaging, but further examination of different oestrous stages using these methods, in addition to examination of programmed cell death (for example using cleaved caspase 3 staining) will further elucidate the mechanisms involved in this turnover.

7.3 Overall conclusions

We posit that the MaSCs/progenitors that generate the ductal network during puberty are distinct unipotent cells that are distributed throughout the gland during ductal morphogenesis. These lineage restricted cells emerge from late embryonic development, demonstrated by marking cells during embryogenesis and tracking their fate into puberty (Figure 7.2). These data are consistent with previously reported work showing unipotent MaSCs being primarily present in the TEBs during puberty, where they proliferate and move towards the subtending duct as it elongates. This is supported by data shown herein, demonstrating that ductal cells also remain proliferative during ductal development. After completion of pubertal morphogenesis, these slow-cycling unipotent MaSCs/progenitors may then be deposited throughout the ductal network for later recruitment in response to pregnancy hormones to generate alveoli in the lactating gland, and even to a lesser extent with oestrous cycling.

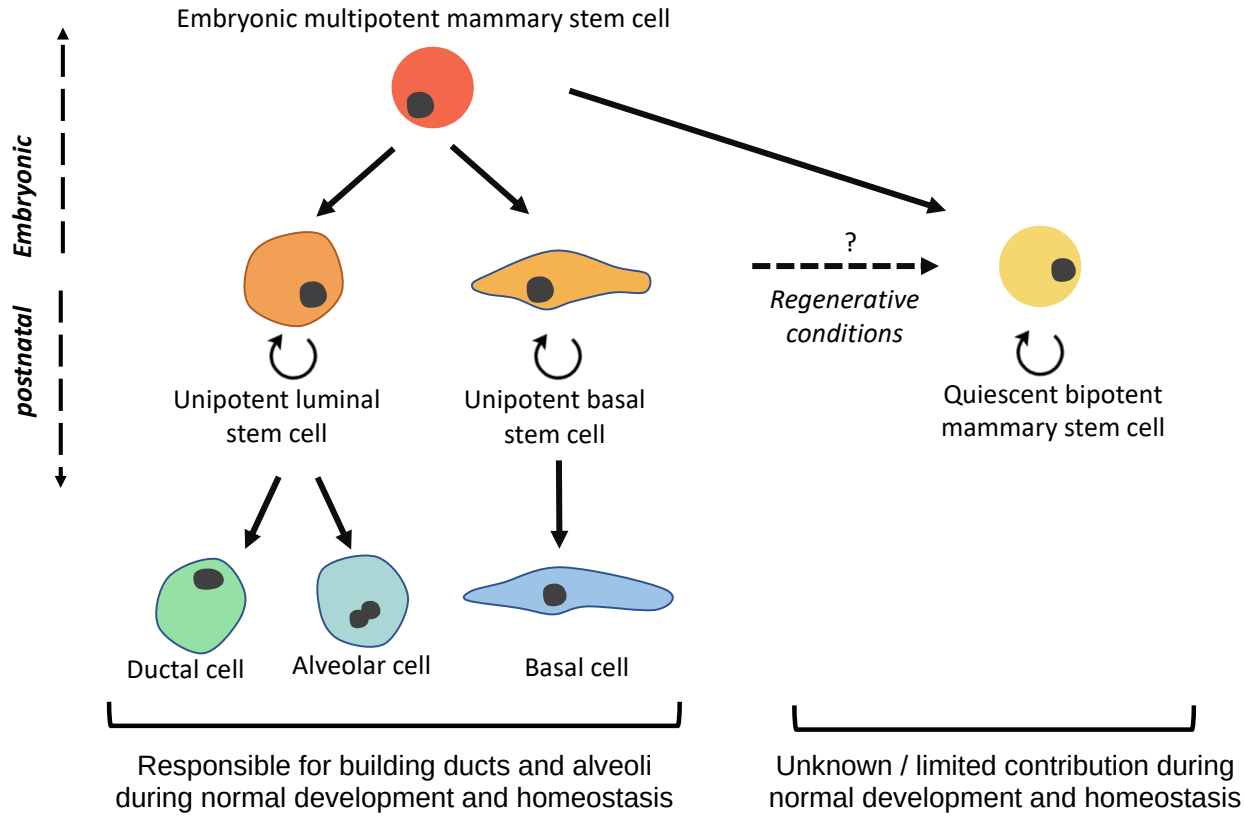


Figure 7.2 Working model of the mammary epithelial cell hierarchy

Multipotent MaSCs are present in the embryo, but some also become lineage restricted during embryonic development. Postnatal mammary gland development (i.e. ductal and alveolar morphogenesis) is principally driven by unipotent luminal and basal mammary stem cells (MaSCs). Luminal stem cells give rise to ductal and alveolar cells that can be ER-positive or -negative. The extent of sub-lineage diversity in the basal compartment, and whether there are distinct ductal and alveolar basal cells, is not yet clear. In addition to the cells responsible for building mammary ducts and alveoli under physiological conditions (left panel), various studies indicate that quiescent and potential stem cells may also reside within the adult mammary gland (right panel). Quiescent bi/multipotent MaSCs (not detected by quantitative or single-cell lineage-tracing approaches) may remain in the mammary gland after embryonic development. In addition, a plastic, intermediate cell type with properties similar to the basal cell lineage may be capable of reverting to a multipotent state under regenerative conditions.

Appendix

Genotyping protocols

Confetti genotyping

	x1 (µL)
10x buffer	2
DNTPs (10 µm)	0.5
Forward primer	1
Reverse primer	1
H20	13.4
Taq	0.1
DNA	2

Primers

Common reverse: 5'-CCAGATGACTACCTATCCTC

Rosa26 forward: 5'-AAAGTCGCTCTGAGTTGTTAT

Confetti mutant forward: 5'-GAATTAATTCCGGTATAACTTCG

Cre forward: 5'-GCGGTCTGGCAGTAAAAACTATC

Cre reverse: 5'-GTGAAACAGCATTGCTGTCACTT

Cycling conditions

94° C	3 min	
94° C	30 sec	
60° C	30 sec	X 35
72° C	1 min	
72° C	5 min	
4° C	∞	

CA30 genotyping

Component	x1 (µL)
10x buffer	2
DNTPs	0.5
Forward primer	1
Reverse primer	1
H20	14.4
Taq	0.1
DNA	1

Primers

Rosa26 forward: ACAGAGGAGCCATAACTGCA

Rosa26 reverse: TGGCAGGAAGGAGGCTTTAA

CA30YFP reverse: GAACTTCAGGGTCAGCTTGC

SYNbglA reverse: CAGGCTCACATCCTCCTCAA.

Cycling conditions

95° C	3 min	
95° C	30 sec	
59° C	1 min	X 35
72° C	2 min	
72° C	5 min	
4° C	∞	

CUBIC reagent recipes

Reagent 1(A)

Ingredient	Sigma cat #	R1 %w/v	R1a %w/v	for 500ml R1a
Urea	U5378	25	10	50g
N,N,N',N'-tetrakis	122262	25	5	25g
Triton x 100	BP151	15	10	50g
5M NaCl	-	-	1:200	2.5mL
dH_20	-	-	-	Up to 500mL

Reagent 2

Ingredient	SIGMA CAT #	ORIGINAL VOL.	FOR 18G
Sucrose	-	50% w/v	8
Urea	U5378	25% w/v	4
Triethanolamine	90279	10% w/v	1.6
Triton x 100	BP151	0.1% w/v	180uL of 10%
dH_20	-	-	To 18g

Mammary organoid media

Basal media

Stock Ingredient	Supplier	50 ML
DMEM/F12	-	47
Glutamax (x100)	Invitrogen, 35050-038	0.5
Hepes (1m)	Invitrogen 15630-056	0.5
Pen/strep (x100)	Invitrogen, 15140122	0.5
N2 supplement	Gibco, 17502	0.5
B27 supplement (w/out vit A)	Gibco, 12587	1.0

Complete media

Ingredient	Supplier	Final conc.	In 1ML
Stock medium	N/A	N/A	980
Neuregulin	R&D, 5898-NR-050	10uL/mL	10 (of 1/10 diluted stock)
Noggin	Peprotech, 250-38	10uL/mL	10 (of 1/20 diluted stock)
R-spondin-1	R&D, 3474-RS-050	100ng/mL	1

Publications arising from this work

Davis, F.M.*, Lloyd-Lewis, B.*, **Harris, O.B**., Kozar, S., Winton, D.J., Muresan, L., et al. (2016). Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat. Commun. 7*: 13053.

Lloyd-Lewis, B.*, Davis, F.M.*, **Harris, O.B**., Hitchcock, J.R., Lourenco, F.C., Pasche, M., et al. (2016). Imaging the mammary gland and mammary tumours in 3D: optical tissue clearing and immunofluorescence methods. *Breast Cancer Res. 18*: 127.

Lloyd-Lewis, B., **Harris, O.B.,** Watson, C.J., and Davis, F.M. (2017). Mammary Stem Cells: Premise, Properties, and Perspectives. *Trends Cell Biol. 27*: 556–567.

Lloyd-Lewis, B.*, Davis, F.M.*, **Harris, O.B**., Hitchcock, J.R., and Watson, C.J. (2018). Neutral lineage tracing of proliferative embryonic and adult mammary stem/progenitor cells. *Development 145*: dev164079.

* *co-first authors*

nature COMMUNICATIONS

ARTICLE

Received 19 Jul 2016 | Accepted 30 Aug 2016 | Published 25 Oct 2016

DOI: 10.1038/ncomms13053 OPEN

Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny

Felicity M. Davis[1,2,*], Bethan Lloyd-Lewis[1,*], Olivia B. Harris[1,3], Sarah Kozar[4], Douglas J. Winton[4], Leila Muresan[5] & Christine J. Watson[1,3]

The mammary gland undergoes cycles of growth and regeneration throughout reproductive life, a process that requires mammary stem cells (MaSCs). Whilst recent genetic fate-mapping studies using lineage-specific promoters have provided valuable insights into the mammary epithelial hierarchy, the true differentiation potential of adult MaSCs remains unclear. To address this, herein we utilize a stochastic genetic-labelling strategy to indelibly mark a single cell and its progeny *in situ*, combined with tissue clearing and 3D imaging. Using this approach, clones arising from a single parent cell could be visualized in their entirety. We reveal that clonal progeny contribute exclusively to either luminal or basal lineages and are distributed sporadically to branching ducts or alveoli. Quantitative analyses suggest that pools of unipotent stem/progenitor cells contribute to adult mammary gland development. Our results highlight the utility of tracing a single cell and reveal that progeny of a single proliferative MaSC/progenitor are dispersed throughout the epithelium.

[1] Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK. [2] School of Pharmacy, The University of Queensland, Brisbane 4072, Australia. [3] WellcomeTrust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 1QR, UK. [4] Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, UK. [5] Cambridge Advanced Imaging Centre, University of Cambridge, Cambridge CB2 1QP, UK. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to C.J.W. (email: cjw53@cam.ac.uk).

The mammary epithelium is comprised of a highly branched, bilayered ductal tree with an inner layer of cytokeratin (K)8/18-expressing luminal cells and a surrounding layer of basal cells that typically express K5/14 and the contractile protein alpha-smooth muscle actin (SMA)[1]. As presumptive targets for transformation in breast cancer, the identity of adult mammary stem cells (MaSCs) and the origin of luminal and basal cell lineages have been the subject of intense investigation and debate[2]. Whilst results from transplantation assays in the mouse mammary gland point to the existence of bi/multipotent MaSCs that reside in the basal compartment[3–5], early genetic fate-mapping studies using lineage-specific promoters demonstrated that distinct unipotent MaSCs maintain the luminal and basal lineages postnatally under physiological conditions[6]. Subsequent lineage-tracing studies have provided further evidence in support of each model[7–10]. However, transplantation assays have been suggested to re-programme cells[6,8,10], and conventional lineage-tracing approaches have relied on prior assumptions regarding the specificity and consistency of expression of lineage markers. Moreover, these studies have induced labelling at levels significantly higher than clonal density, thus confounding analysis of daughter cells and their contribution to different lineages. Consequently, we have sought to resolve these complexities and determine unequivocally the potential of adult MaSCs, during puberty and pregnancy, by combining the use of mouse models that enable labelling of a single random cell and all of its progeny *in situ* with confocal three-dimensional (3D) imaging.

Results

Optical clearing and 3D imaging of the intact mammary gland. To accurately determine the capacity of a single marked stem or progenitor cell and its progeny to contribute to the development of the branching mammary epithelial network *in vivo*, the entire ductal tree needs to be visualized at high spatial resolution. The utility of 3D imaging for fate-mapping studies has previously been demonstrated in stroma-divested mammary glands[7,9]. In this study we developed and refined methods for 3D imaging in the mammary gland, using techniques for optical tissue clearing to enable visualization of the mammary epithelium at single-cell resolution, without the need for enzymatic digestion or mechanical dissection.

Tissue clearing methods that have been developed are based on mitigating light scattering caused by cellular and extracellular structures with different refractive indices[11]. However, the utility of these protocols in the mammary gland remains largely unexplored. We determined that the SeeDB[12] and the CUBIC[13] tissue-clearing protocols provide superior optical clarity in mammary tissue (Fig. 1a). Moreover, by combining optical tissue clearing with wholemount immunostaining (Supplementary Fig. 1a) and algorithms to improve the signal-to-noise ratio of 3D image sequences[14] (Supplementary Fig. 1b), we were able to visualize the epithelial ductal tree to depths exceeding 400 μm, identifying K5-expressing basal cells and K8-expressing luminal cells at single-cell resolution (Fig. 1b, Supplementary Fig. 2 and Supplementary Movies 1 and 2). We could also image lactating mammary glands in 3D, highlighting the organization of basal cells by SMA immunostaining and luminal cells by their expression of E-cadherin and K8 (Fig. 1c and Supplementary Movie 3). We noted that K8 is non-uniformly expressed in luminal cells (Fig. 1b,c and Supplementary Movie 2) and identified two distinct subpopulations comprising $K8^{lo}$ and $K8^{hi}$ cells *in situ*, the latter co-staining with nuclear progesterone receptor (PR) (Fig. 1d and Supplementary Fig. 3). While these K8 subpopulations are present in similar proportions in virgin ducts, $K8^{hi}$ cells are extremely sparse in lactational alveoli (Fig. 1c and Supplementary Fig. 3c). An association between high K8/18 expression and the functionally distinct $CD24^{hi}$/prominin-1^{+}/$Sca1^{+}$ hormone-sensing luminal population has previously been observed[15]. Moreover, lineage-tracing studies using K8-$CreER^{T2}$/Tomato-reporter mice have demonstrated preferential genetic labelling of $CD24^{hi}$/$Sca1^{+}$ luminal cells[16]. Collectively, these observations question the utility of the promoters of keratins and other presumed lineage-specific genes as suitable drivers of reporter proteins for lineage-tracing studies.

A stochastic labelling strategy for single-cell lineage tracing. To avoid prior assumptions regarding the expression profile of MaSCs and to track the fate of a single marked cell in the mammary epithelium, we utilized $R26^{[CA]30}$ reporter mice, which have previously been used to infer stem cell dynamics in the intestinal epithelium[17]. This model encompasses a dinucleotide repeat tract, $[CA]_{30}$, positioned downstream of the translational start site of an out-of-frame reporter gene (enhanced yellow fluorescent protein (EYFP) or modified β-glucosidase (SYNbglA)) inserted in the constitutively expressed Rosa26 locus (Fig. 2a). The inherent instability of microsatellite repeats can lead to spontaneous, random frame-shift mutations during DNA replication, which may place the reporter gene in-frame, thereby resulting in its expression. The advantages of this labelling approach are twofold: first, replication slippage is equally likely to occur in all cycling cells; and second, strand slippage is extremely rare[17], thus allowing all of the progeny of a single labelled cell to be identified with confidence.

Clonal labelling patterns in the mouse mammary gland. To determine the suitability of this model for single-cell lineage tracing in the mammary epithelium we examined clone abundance, size and distribution in $R26^{[CA]30SYNbglA}$ mice during pubertal development, when functionally active MaSCs are presumed to drive ductal elongation and branching morphogenesis[18,19]. These mice contain a modified β-glucosidase gene, which is thermostable and resistant to epigenetic silencing, downstream of the $[CA]_{30}$ tract (Fig. 2a), enabling macroscopic clonal analysis by wholemount histochemistry. Using this model, combined with CUBIC-based tissue clearing, regions of ducts containing variable numbers of β-glucosidase^{+} cells interspersed with unlabelled cells could be visualized *in situ* (Fig. 2b–d and Supplementary Figs 4 and 5). As in the intestine, strand slippage was extremely rare in the mammary epithelium, with $\sim 1.49 \pm 0.92$ total labelling events observed per gland (Supplementary Fig. 4) and, as such, the likelihood of clone convergence in this model is exceedingly low. We observed large contiguous clonal regions containing several hundred label-positive cells that spanned numerous branching ducts (Fig. 2b,c and Supplementary Fig. 5). These were considered to have arisen from a single MaSC or progenitor. Isolated regions that contained limited numbers of label-positive cells were also observed (Supplementary Fig. 4), most likely the result of strand slippage in more differentiated cells or in progenitors with restricted replicative potential (for example, *Elf5*-expressing luminal progenitors[7]). However, given the continual, albeit rare, genesis of labelled cells in this model, the possibility of recent strand slippage in a MaSC or highly proliferative progenitor could not be excluded. Label-positive regions were also detected after multiple pregnancy/involution cycles (Supplementary Fig. 6), indicating that some progeny may be long-lived.

Whilst strand slippage in a germ cell resulted in complete and uniform genetic labelling of ducts (Supplementary Fig. 7), clonal

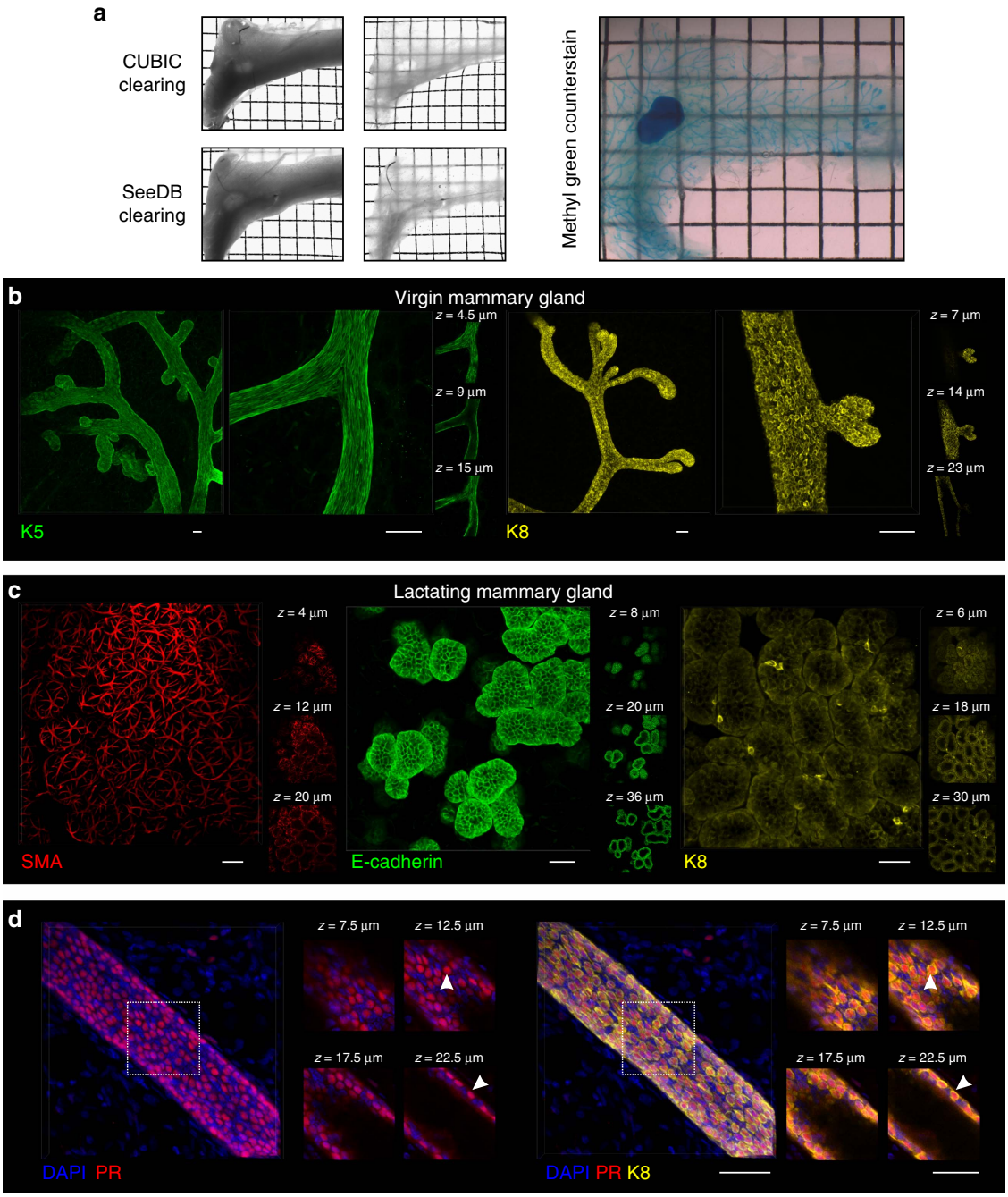
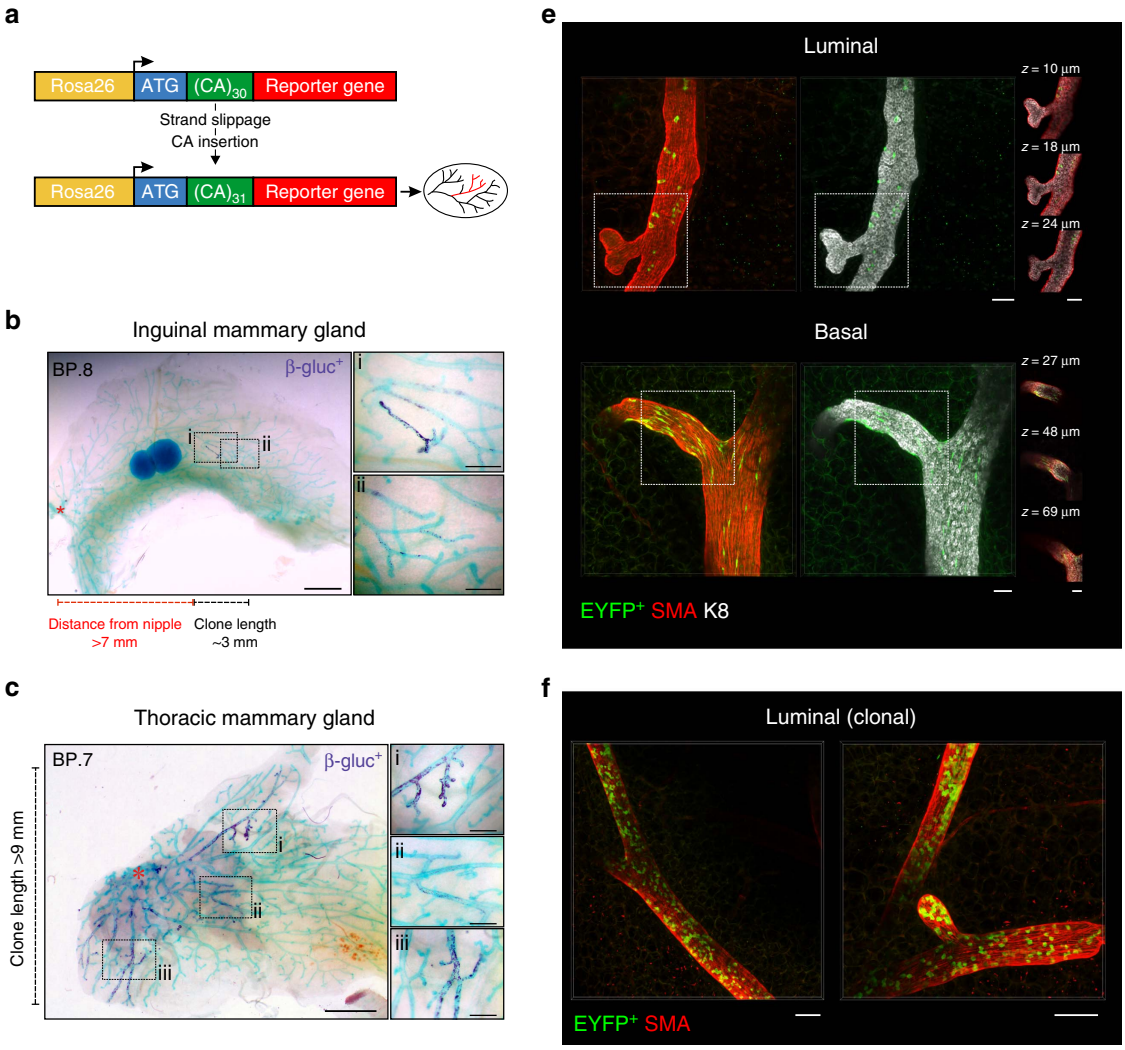


Figure 1 | Optical tissue clearing and 3D imaging of the mammary gland. (**a**) Transmission images of an entire virgin mammary gland before (left panels) and after (middle panels) tissue clearing using CUBIC or SeeDB, and a CUBIC-cleared mammary gland counterstained with methyl green to visualize the complete ductal network (right panel). Representative images from three mice. Grid width: 2 mm. (**b**) 3D imaging of K5 and K8 immunostaining of SeeDB-cleared virgin mammary tissue within its native stroma. K5 overview shows 1.18 mm (*xy*) of mammary gland ($z = 114$ µm imaging stack depth); K8 overview shows 834 µm (*xy*) ($z = 114$ µm). The depth (*z*) is relative to the first image in the image sequence, reached after passing through the mammary fat pad (~350 µm). Scale bars, 50 µm. (**c**) Immunostaining for SMA ($xy = 579$ µm; $z = 53$ µm), E-cadherin ($xy = 467$ µm; $z = 36$ µm) and K8 ($xy = 399$ µm; $z = 32$ µm) in mammary glands from lactating mice. Scale bars, 50 µm. (**d**) Immunostaining shows populations of $K8^{hi}$ and $K8^{lo}$ luminal cells, with the $K8^{hi}$ cells costaining with nuclear PR (arrowhead) (representative images from three mice); scale bars, 50 µm.

expansion from a single MaSC/progenitor produced a stochastic labelling pattern, with β-glucosidase$^+$ cells intermixed randomly with unlabelled cells in branching ducts spanning over 8 mm in length (Fig. 2b–d and Supplementary Fig. 5). These clonal labelling patterns strongly suggest that a pool of active mammary stem/progenitor cells reside within each terminal end bud (TEB), the presumptive origin of MaSCs[20], and contribute to the development of each major duct during puberty. The unequal distribution of labelled progeny between branching ducts (for example, Fig. 2b) is most likely due to the dilution of marked daughter cells with the progeny of unmarked MaSCs/progenitors during TEB bifurcation or secondary branching.

Labelled progeny arising from the expansion of a single β-glucosidase$^+$ cell had a luminal-like morphology (Fig. 2b and



d

R26$^{(CA)30SYNbglA}$								
Total no.	30 mice (240 mammary glands)							
MaSC clone ID	BP.1	BP.2	BP.3	BP.4	BP.5	BP.6	BP.7	BP.8
Linear length (mm)	>8	~5	>4	>3	~3	~4	>9	~3
No. branches	>18	>10	>15	>12	>19	>9	>20	>7
Distance from origin (mm)	>3	>1	>5	>1	>1	>3	0	>4
Shown in (Fig ref.)	S5	S5	S5	S5	S5	S5	2c	2b

g

R26$^{(CA)30EYFP}$			
Total no.	63 mice (504 mammary glands)		
MaSC clone ID	YP.1	YP.2	YP.3
Linear length (mm)	>8	>7	>4
No. branches	>20	>10	>9
Distance from origin (mm)	>2	>2	>2
Composition (Lu/Ba/Mixed)	Lu	Lu	Ba
Shown in (Fig ref.)	3, S9	S8	4

Figure 2 | Single-cell lineage tracing in the virgin mammary gland. (**a**) Schematic representation of the *R26*$^{[CA]30}$ mouse model. (**b**,**c**) Examples of two large clonally marked regions (BP.8 and BP.7) in mammary glands from *R26*$^{[CA]30SYNbglA}$ mice that were likely to have arisen from the labelling of a MaSC/progenitor (based on linear length and number of label-positive branches) (**d**). Dark purple staining is β-glucosidase$^+$ cells; mammary tissue was counterstained with methyl green. Annotations show the linear length of the clones and their distance from the nipple region (asterisk). Clone BP.7 originated in the nipple region and is likely to have been labelled very early in development. Scale bars, 2 mm (overview) and 0.5 mm (inset). (**d**) A summary of the eight clonally marked regions likely to have arisen from the labelling of a MaSC/progenitor, observed from the analysis of 30 *R26*$^{[CA]30SYNbglA}$ mice. (**e**) Examples of luminal (top panel) and basal (bottom panel) EYFP$^+$ cells from *R26*$^{[CA]30EYFP}$ mice representing over 25 label-positive regions. Scale bars, 50 µm. (**f**) A large clonally marked region containing many EYFP$^+$ cells. Labelled progeny spanned multiple ducts and exhibited a sporadic labelling pattern, intermixed with unlabelled cells. Scale bars, 100 µm. (**g**) A summary of three clonally marked regions presumed to have arisen from the labelling of a MaSC/progenitor, observed from the analysis of 63 *R26*$^{[CA]30EYFP}$ mice. Lu, luminal; Ba, basal.

Supplementary Fig. 5) in the majority of label-positive regions. Regions with basal-like β-glucosidase$^+$ cells were observed but were less common, possibly reflecting a smaller proportion of proliferating cells in this cellular compartment[21,22]. On one occasion we observed ducts that appeared, on the basis of morphology, to be comprised exclusively of β-glucosidase$^+$

luminal cells connected to ducts comprised exclusively of β-glucosidase$^+$ basal cells (Fig. 2c(i) and (iii) versus (ii)). The proximity of these diverging ducts suggests that labelled cells arose from a single, bipotent parent cell that gave rise to one luminal and one basal daughter. Given the expansive nature of this exceptional clone (BP.7, >9 mm in length), and the presence of labelled cells in the nipple region, it is likely that a bipotent MaSC was genetically marked very early in development of the gland, possibly during embryogenesis[6,23], and later gave rise to lineage-restricted progeny in the postnatal gland. This juxtaposition of presumptive luminal and basal ductal clones provides intriguing new insights into the likely fate of progeny of a bipotent embryonic MaSC that could only be revealed by single-cell labelling.

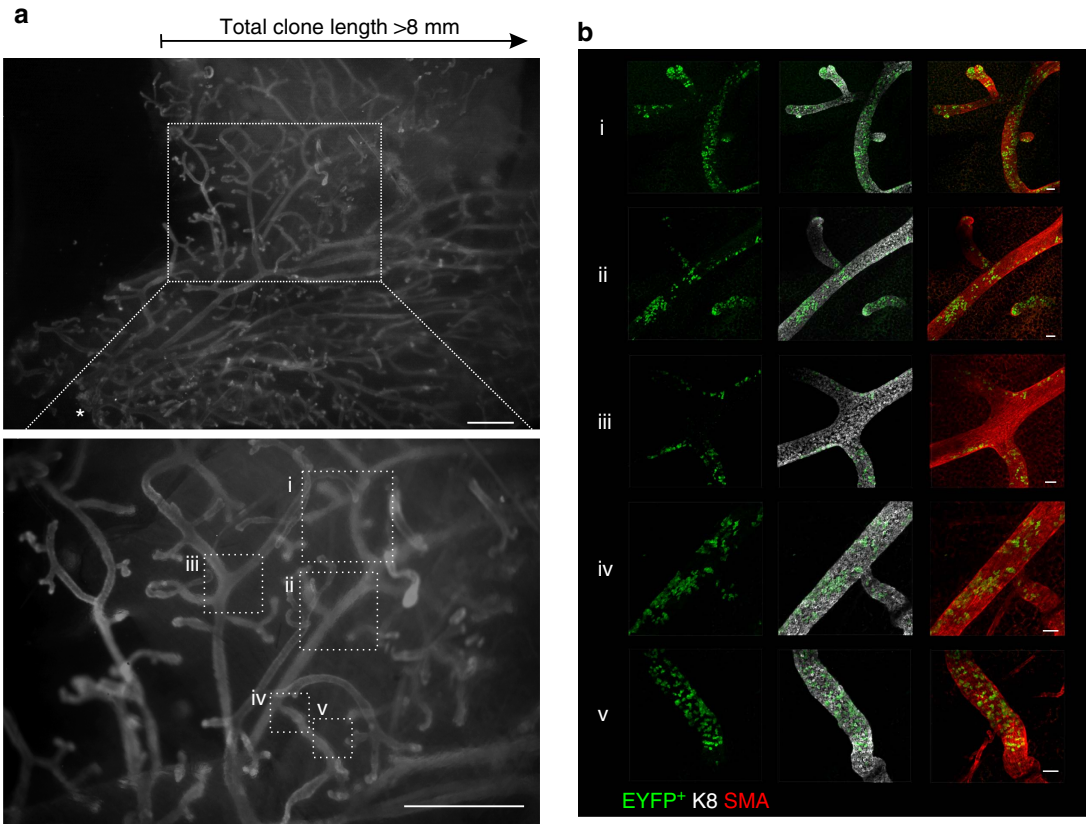
Unipotent cells contribute extensively to ductal morphogenesis. To examine more closely the clonal labelling patterns arising from adult MaSCs/progenitors, and to more definitively determine lineage on the basis of appropriate markers, we utilized the *R26*$^{[CA]30EYFP}$ mouse model, combined with SeeDB-based optical tissue clearing. Using this approach we were able to visualize and characterize progeny arising from a single fluorescently marked cell *in situ* with single-cell resolution (Fig. 2e,f). We note that despite the high degree of optical clarity achieved using this method, some regions deep within the mammary fat pad could not be visualized at single-cell resolution by confocal microscopy and thus a larger number of mice were required for analysis in this model. Immunolabelling for markers of basal (SMA) and luminal (K8) lineages confirmed that the majority of labelled clones were luminal, with few basal clones observed (Fig. 2e). Only one large EYFP$^+$ basal clone, spanning over nine branches, which could have arisen from a stem cell, was observed (Fig. 2g). Clonal expansion of a single EYFP$^+$ luminal cell produced a mosaic labelling pattern identical to those observed in the *R26*$^{[CA]30SYNbglA}$ model (Fig. 2f), confirming that more than one luminal MaSC/progenitor contributes to the elongation of each major duct during puberty. Since the timing of the slippage event cannot be determined, we measured both the length of each clone and the distance from the nipple region where labelled cells are first observed (Fig. 2g and Supplementary Fig. 8, clone YP.2). If the labelled cell of origin is more than 1 mm from the nipple region, our assumption is that slippage has occurred in a stem or progenitor cell postnatally[20]. All such large clonal regions were lineage-restricted and we did not detect luminal and basal EYFP$^+$ cells intermingled within the same duct (Fig. 2g). These data support our observations with the *R26*$^{[CA]30SYNbglA}$ mice and provide further compelling evidence that unipotent MaSCs/progenitors contribute extensively to ductal morphogenesis[6]. However, due to the requirement for proliferation to label and trace stem and progenitor cells and the aforementioned issues associated with deep imaging, we cannot rule out the possibility that rare quiescent bipotent MaSCs, not detected in the *R26*$^{[CA]30EYFP}$ model, may exist.

To confirm the lineage restriction of adult MaSCs/progenitors, and quantify their contribution to ductal morphogenesis, we analysed a large clonal region (Fig. 3a–e and Supplementary Fig. 9) using imaging algorithms for the volumetric segmentation of mammary ducts and the subsequent characterization of all ductal EYFP$^+$ cells (Methods and Fig. 3c). This clonal region was more than 8 mm in length and comprised over 20 branching ducts (Figs 2g and 3a and Supplementary Fig. 9). Since the clone originated more than 2 mm from the nipple region, this suggests that strand slippage most likely occurred postnatally. All EYFP$^+$ cells examined by 3D analysis expressed the luminal marker K8 (Fig. 3b), and encompassed both K8^{hi} and K8^{lo} subpopulations (Fig. 3c and Supplementary Movie 4), with a modest but significant overrepresentation of the EYFP label in K8^{hi} cells (Kolmogorov–Smirnov test at $P<0.05$) (Supplementary Fig. 10a). The lineage restriction of this clone was also confirmed by histological analysis of sectioned tissue over depths of 300 μm (Fig. 3d and Supplementary Fig. 10b). To determine the potential contribution of an active MaSC/progenitor to each duct, a volumetric ratio of EYFP$^+$ cells with respect to the total cellular volume was computed, revealing that the parent cell contributed on average 4.7 ± 1.7% of the total cellular volume in this region of the clone (Fig. 3e). A similar analysis was performed for a basal clone (Fig. 4a–c), where it was determined that the parent basal MaSC/progenitor contributed on average 5.8 ± 3.2% of the total cellular volume (Fig. 4d). Whilst these numbers may reflect the differential proliferative and competitive behaviours of stem/progenitor cells and their progeny, in addition to their random distribution with branching, these data indicate that there may be at least 20 luminal and 15 basal lineage-restricted stem/progenitor cells in each major duct driving mammary gland morphogenesis during puberty.

Neutral lineage tracing using a multicolour reporter. Although the *R26*$^{[CA]30SYNbglA}$ and *R26*$^{[CA]30EYFP}$ models have provided unprecedented insights into the contribution of a single stem/progenitor cell to mammary gland development, we sought to confirm the labelling pattern during pubertal development with a different neutral approach that also allows the timing of the labelling event to be controlled and is not dependent on a cell being in cycle at the time of labelling. To achieve this, we utilized the Confetti multicolour reporter mouse[24], combined with wholemount immunostaining and 3D imaging. We generated mice that were hemizygous for both *R26-Confetti* (ref. 24) and *R26-CreERT2* (ref. 25; Fig. 5a) and administered a single low-dose of tamoxifen to 4-week-old mice followed by a 3-week chase to trace the progeny of cells labelled at the onset of puberty (Supplementary Fig. 11a). This resulted in low-frequency[26] multicolour labelling in the mammary epithelium, allowing us to distinguish individual clones. To identify luminal and basal cells, wholemount immunostaining was performed using a lineage-specific marker (either K8 or SMA, respectively) and tissues were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) to localize the mammary epithelium. Cyan fluorescent protein (CFP)-expressing clones were under-represented and were therefore not analysed. The explanation for this is not entirely clear but may relate to the poor penetration of short-wavelength light through thick specimens and the fine membranous localization of the CFP reporter protein. Using this approach, we were able to visualize luminal and basal lineage-restricted GFP$^+$, YFP$^+$ and RFP$^+$ clones (Fig. 5b,c and Supplementary Fig. 11b,c). Notably, labelling patterns in the confetti mice were similar to those in the *R26*^{CA30} models, validating the latter approach.

Unipotent cells contribute extensively to alveologenesis. Finally, to investigate the contribution of a single adult MaSC/progenitor to the formation of lobuloalveolar structures during pregnancy, we analysed clonal labelling patterns in lactating *R26*$^{[CA]30EYFP}$ mice (Fig. 6 and Supplementary Figs 12 and 13). We noted an increased number of clonal regions in tissue from lactating mice compared with puberty (probably as a consequence of the higher levels of proliferation), and a striking variety of patterns. The unequal distribution of EYFP$^+$ cells between lobuloalveolar units could suggest that an alveolar stem cell niche is situated close to the branch point of the subtending ducts (Fig. 6a). Competition for niche occupancy may dictate the dispersal of labelled and

unlabelled daughter cells between adjacent lobules, with further competitive interactions between their respective progeny possibly determining labelling outcomes within each alveolus[27]. Indeed, alveoli that were comprised almost entirely of label-positive luminal cells were occasionally observed (Fig. 6b), implying that the descendants of any unmarked luminal stem/progenitor cell had been outcompeted. We also observed an interesting pattern where many alveoli within a lobule contained only a single EYFP$^+$ cell (Fig. 6c) suggesting the possibility that the original labelled cell is restricted to a lineage that constitutes a



d

	EYFP$^+$ cells (2D analysis)	EYFP$^+$ events (3D analysis)
Luminal	281/281 (100%)	227/227 (100%)
Basal	0/281 (0%)	0/227 (0%)

e

Image sequence	Ductal EYFP$^+$ cells/total cellular volume (%)
1	4.08
2	2.55
3	6.38
4	5.66
Avg±S.D	4.7±1.7

minor, although possibly important, component of each alveolus. This has not been observed before and the identity and function of these cells is unclear. Similar to the pubertal epithelium, all large clonally marked regions were lineage-restricted, with separate luminal and basal clones observed during lactation (Fig. 6d and Supplementary Fig. 13). Of note, luminal cells contributed to both the $K8^{hi}$ hormone-sensing lineage, as well as the more abundant $K8^{lo}$ alveolar lineage (Supplementary Fig. 13, clone YL.1). In all cases, $EYFP^+$ progeny was intermixed with unlabelled cells in a polyclonal pattern that spanned numerous lobuloalveolar structures (Fig. 6d). Quantification of labelling patterns in these large clonally marked regions, some comprising over 100 alveoli (Fig. 6e), revealed that the majority of alveoli were comprised of both $EYFP^+$ cells and unlabelled cells of a single lineage (Fig. 6f), demonstrating that most alveoli are derived from at least two lineage-restricted stem/progenitor cells. These data contrast with previous studies suggesting that alveoli can be comprised of the progeny of a single cell[7,8]. The reason for this discrepancy is not clear but possibly reflects the uncertainties in analysing progeny of co-incident clones. Alternatively, stochastic stem cell fate could result in neutral drift and elimination of other stem cells and their progeny[28]. These insights into alveolar stem cell biology reflect the power of *in vivo* lineage tracing at clonal density.

Discussion

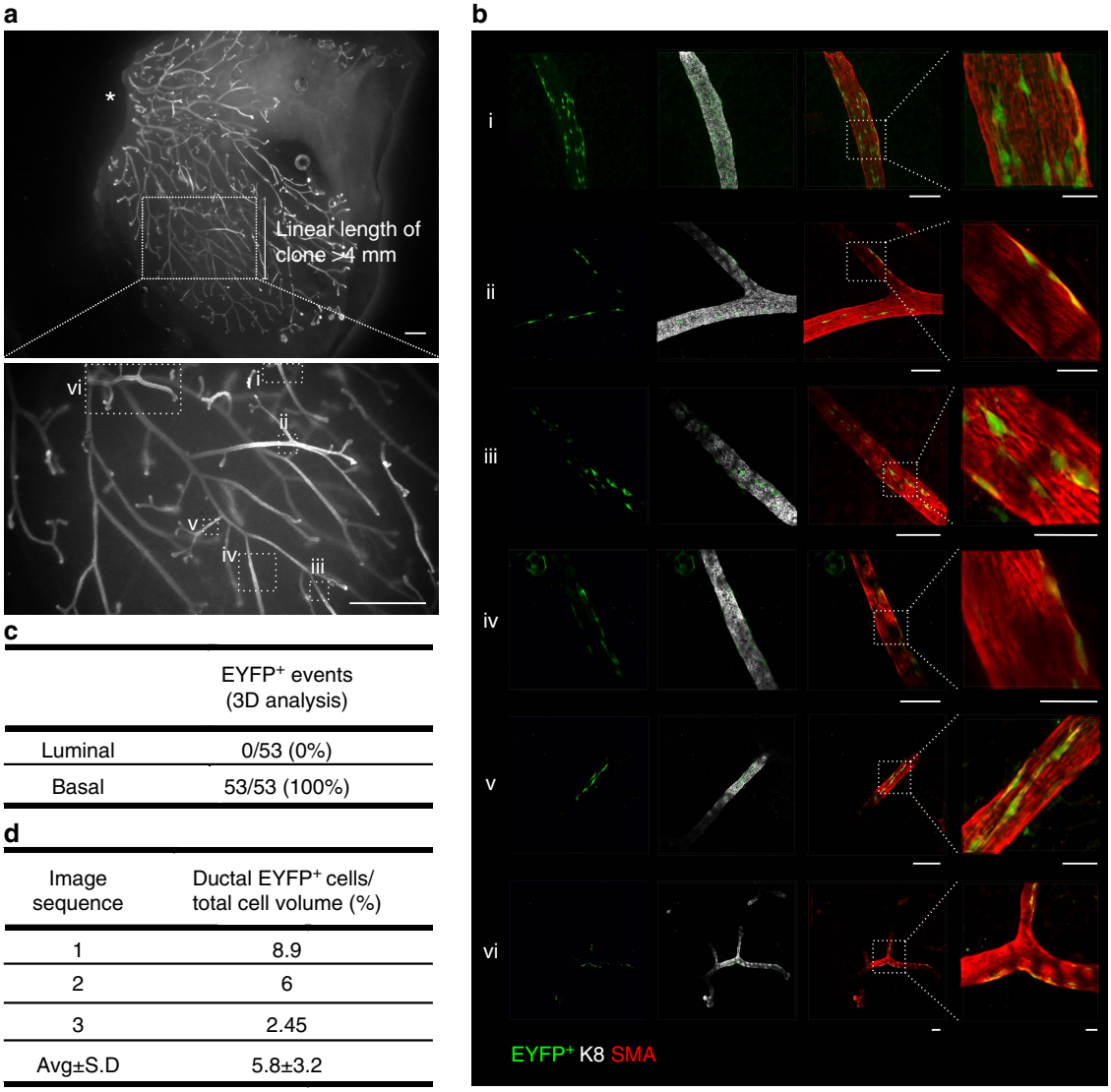
The existence of MaSCs was demonstrated over 50 years ago[29]. More recently, the identity and potential of these cells has come under intense scrutiny, yet a number of uncertainties remain. Prime amongst these is whether MaSCs in the adult are unipotent or bipotent[6–9]. Although this may appear to be a relatively straightforward question to address, current experimental approaches have not provided an unequivocal answer. This is primarily a consequence of their dependence on presumed lineage-restricted promoters to drive reporter gene expression in a significant proportion of MaSCs for population-based fate tracking. Using this approach the probability of two or more clones arising in the same region is high, confounding their analysis. Thus, promiscuous labelling and subsequent expansion of even a single lineage-restricted MaSC could resemble clonal expansion of a bipotent MaSC[26]. In this context it is important to note that the expression of K14 and K18 is differentially regulated in the pre-pubertal mammary gland with some luminal cells unexpectedly expressing K14 (refs 30,31). Similar difficulties arise with the K8 promoter that is expressed at low levels in a subset of cells, leading to the disparate labelling of luminal cell populations in fate-mapping studies[16]. Thus, lineage tracing with these promoters is not definitive for the assessment of potency. We therefore adopted two agnostic fate-mapping strategies to avoid these confounding issues and this has resulted in a number of unanticipated observations and valuable insights that could only have been revealed by this stochastic single-cell-labelling approach.

Our first intriguing observation was the random distribution of labelled progeny of a single cell to multiple ducts (depicted schematically in Fig. 7a). Some regions had a high density of labelled cells while others had a much lower density, indicating the presence of multiple lineage-restricted stem cells and the admixing of their progeny. Imaging of entire mammary glands also revealed that all labelled ducts were connected, suggesting bifurcation and branching from a TEB in which the labelled cell presumably arose during puberty. We investigated the nature of cells within large clones using 3D imaging algorithms, and revealed that all labelled cells within these regions were lineage restricted. Furthermore, based on a volumetric analysis and the assumption that all MaSCs/progenitors have the same capacity to contribute to ductal outgrowth, we estimate that at least 20 luminal and 15 basal MaSCs/progenitors contribute to the growth of a major duct. By extension, this would equate to a few hundred unipotent luminal and basal MaSCs/progenitors per gland, which drive ductal morphogenesis during puberty. Our lactation data reveal the unexpected presence of different subpopulations of alveolar cells, including one type that contributes only a single cell to most alveoli in a lobuloalveolar cluster. We suggest that, as alveolar expansion during pregnancy is critically important, and may occur several times in a lifetime, a pool of alveolar stem or committed progenitor cells is required. The variable contribution of cells to individual alveoli (depicted schematically in Fig. 7b), with 100% contribution being rare, could reflect prior commitment to specific lineages or competition for the stem/progenitor cell niche[32,33].

We posit that the MaSCs/progenitors that generate the ductal network during puberty are distinct from those that have a more homeostatic function in the adult, the latter possibly arising from bipotent embryonic MaSCs that may persist after birth and remain quiescent[30]. Indeed, these cells would not have been labelled by our approach. Nevertheless, our data are consistent with unipotent mammary stem/progenitor cells being primarily present in the TEBs during puberty, where they proliferate and move towards the subtending duct as it elongates. These TEB-resident MaSCs/progenitors would be lost when the TEBs regress at the completion of puberty. However, slow-cycling unipotent MaSCs/progenitors may be deposited throughout the ductal network[34,35] and could later be recruited in response to pregnancy hormones to generate alveoli.

Our work has illuminated the capacity of a single cell in the adult mammary gland to contribute to mammary gland development. A complete resolution of the mammary stem cell hierarchy controversy will require the ability to label a single cell at a defined moment and follow its fate over time. Although fraught with difficulties, prospective isolation and transcriptome analysis of single MaSCs will be an aim for the future.

Figure 3 | 3D analysis of a clone arising from a single labelled luminal presumptive MaSC. (**a**) Wholemount fluorescence images (K8 immunofluorescence) of the mammary ductal network demarcating the linear length of the clone and one region (magnified views, i–v) that was imaged at high cellular resolution by confocal microscopy in **b**. Further regions from this clone are shown in Supplementary Fig. 9. Asterisk shows the location of the nipple. Scale bar, 1 mm (wholemount) and 50 μm (confocal). (**c**) Images of a clonally marked region that was analysed by 3D image analysis. Digital segmentation of $EYFP^+$ cells within the luminal (K8-expressing) and basal (SMA-expressing) compartments is shown. Original 3D images show that progeny from a single luminal MaSC/progenitor included both $K8^{hi}$ (arrow) and $K8^{lo}$ (arrowhead) cells. BV, blood vessel. Scale bars, 1 mm (wholemount) and 100 μm (confocal). (**d**) Tabulated results of 3D and two-dimensional (2D) clonal analyses. For 3D analysis, all segmented ductal $EYFP^+$ cells were classified as luminal based on the proportion of K8 versus SMA signal ($n = 227$ cells from 4 image sequences). For 2D analysis, cells were classified by manual scoring of histological sections ($n = 281$ cells from 10 sections spanning 300 μm depth). This clone (YP.1) is one of three clones likely to have arisen from the labelling of a MaSC (based on linear length and number of label-positive branches), identified from the analysis of over 500 mammary glands from 63 hemizygous $R26^{[CA]30EYFP}$ pubertal mice. (**e**) Tabulated results of the computed volumetric ratio of $EYFP^+$ cells with respect to cellular volume for each of the four regions analysed.



c

	EYFP+ events (3D analysis)
Luminal	0/53 (0%)
Basal	53/53 (100%)

d

Image sequence	Ductal EYFP+ cells/ total cell volume (%)
1	8.9
2	6
3	2.45
Avg±S.D	5.8±3.2

Figure 4 | 3D analysis of a clone arising from a single labelled basal presumptive MaSC. (**a**) Wholemount fluorescence images (K8 immunofluorescence) of the mammary ductal network, mapping regions (i–vi) that were imaged using a confocal microscope in **b**. All EYFP$^+$ cells in this clone (YP.3) were basal. A distinct clone (separated by >1 mm) was identified in this tissue piece and contained only luminal cells (shown in Supplementary Fig. 8). Luminal and basal EYFP$^+$ cells were not imaged in the same branches and clones were never intermixed. Asterisk shows the nipple and origin of the ductal network. Scale bar, 1 mm (wholemount) and 100 µm (confocal, overview) or 30 µm (confocal, magnified view). (**c**) Tabulated results of the 3D lineage analysis ($n = 53$ cells from 3 image sequences). (**d**) Tabulated results of the computed volumetric ratio of EYFP$^+$ cells with respect to total basal cellular volume for each of the three regions analysed.

Methods

Animal models. *R26*$^{[CA]30SYNbglA}$ and *R26*$^{[CA]30EYFP}$ mice (on a C57Bl/6J background)[17] were provided by Prof. Douglas Winton (Cancer Research UK Cambridge Institute). Female virgin mice were killed by dislocation of the neck or terminal anaesthesia at 7 weeks of age for all studies in puberty. For single-cell lineage tracing in lactating mice, female *R26*$^{[CA]30}$ mice were mated with C57Bl/6J male studs and tissue was collected between lactation days 2–4. For analysis of multi-parous mice, female *R26*$^{[CA30]SYNbglA}$ mice were mated with C57Bl/6J male studs (for 3 pregnancy/involution cycles), and allowed to naturally litter and wean their pups. The final wean was followed by an 8- to 9-week interval before mammary tissue was collected. All quantitative analyses were performed on mice that were hemizygous for *R26*$^{[CA]30SYNbglA}$ or *R26*$^{[CA]30EYFP}$. Mice that were hemizygous for both *R26-Confetti* (ref. 24) and *R26-CreERT2* (ref. 25) (*R26-Confetti;R26-CreERT2*) were generated by mating homozygous mice. These mice can theoretically produce cells that express either membrane-bound cyan, nuclear green, cytoplasmic yellow or cytoplasmic red fluorescent proteins following tamoxifen administration. Eight mammary glands (excluding the first (cervical) pair) were dissected and analysed for each mouse. Mammary glands were excised and fixed in 10% neutral buffered formalin (NBF) for 9 h at room temperature. Animals were housed in individually ventilated cages under a 12:12 h light–dark cycle, with water and food available *ad libitum*. All animal experimentation was carried out in accordance with the Animal (Scientific Procedures) Act 1986, the European Union Directive 86/609 and with local ethics committee approval. No statistical method was used to predetermine sample size.

Induction of lineage tracing. Lineage tracing was induced at 28 days for studies in puberty in *R26-Confetti;R26-CreERT2* mice. A single intraperitoneal injection of tamoxifen (1 mg) in sunflower oil was administered and tissues were collected after a 2-day chase to determine initial labelling or after 3 weeks for pubertal lineage-tracing studies.

Reagents. The following reagents were purchased from Sigma Aldrich: NBF; urea; *N,N,N′,N′*-tetrakis(2-hydroxypropyl)ethylenediamine; 2,2′,2″-nitrilotriethanol; fructose; α-thioglycerol; DAPI dilactate; 3,3′-diaminobenzidine tetrahydrochloride (DAB); and tamoxifen. Sucrose was purchased from Fisher Scientific. Triton X-100 was purchased from VWR International. The following primary antibodies were used for immunostaining: rabbit anti-K5 (Covance, PRB160P, 1:100); rat anti-K8 (Developmental Studies Hybridoma Bank, TROMA-I, 1:50); rabbit anti-SMA (Abcam, ab5694, 1:300); mouse anti-SMA (Abcam, ab7817, 1;200); rabbit

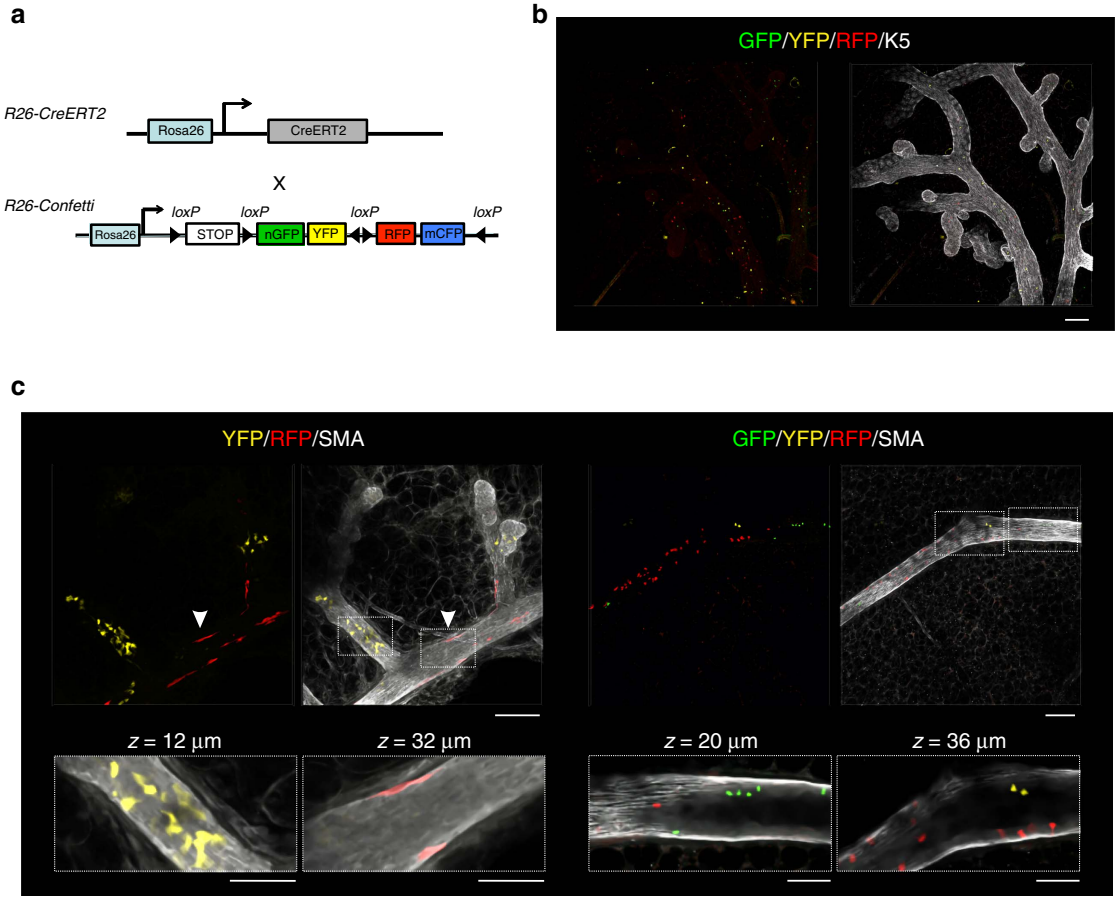


Figure 5 | Clonal labelling patterns observed in *R26-Confetti;R26-CreERT2* pubertal mice. (**a**) Schematic representation of the *R26-Confetti;R26-CreERT2* mouse model. (**b**) Initial labelling level observed 2 days after the administration of a single, low-dose of tamoxifen (1 mg intraperitoneal (i.p.)) to 4-week-old mice. Scale bar, 100 µm. (**c**) Labelling patterns observed in *R26-Confetti;R26-CreERT2* pubertal mice confirm the results from the $R26^{[CA]30}$ model. Labelling was induced by the administration of a single, low-dose of tamoxifen (1 mg i.p.) to 4-week-old mice and mammary glands were collected after a 3-week chase. Left panel shows a region containing YFP^+ luminal cells and RFP^+ basal cells (arrowhead) populating three branches, interspersed with unlabelled cells. Right panel shows a region containing GFP^+, YFP^+ and RFP^+ luminal cells in a single branch. Scale bar, 100 µm (overview) and 50 µm (inset). Images are representative of three mice. Additional images are shown in Supplementary Fig. 11.

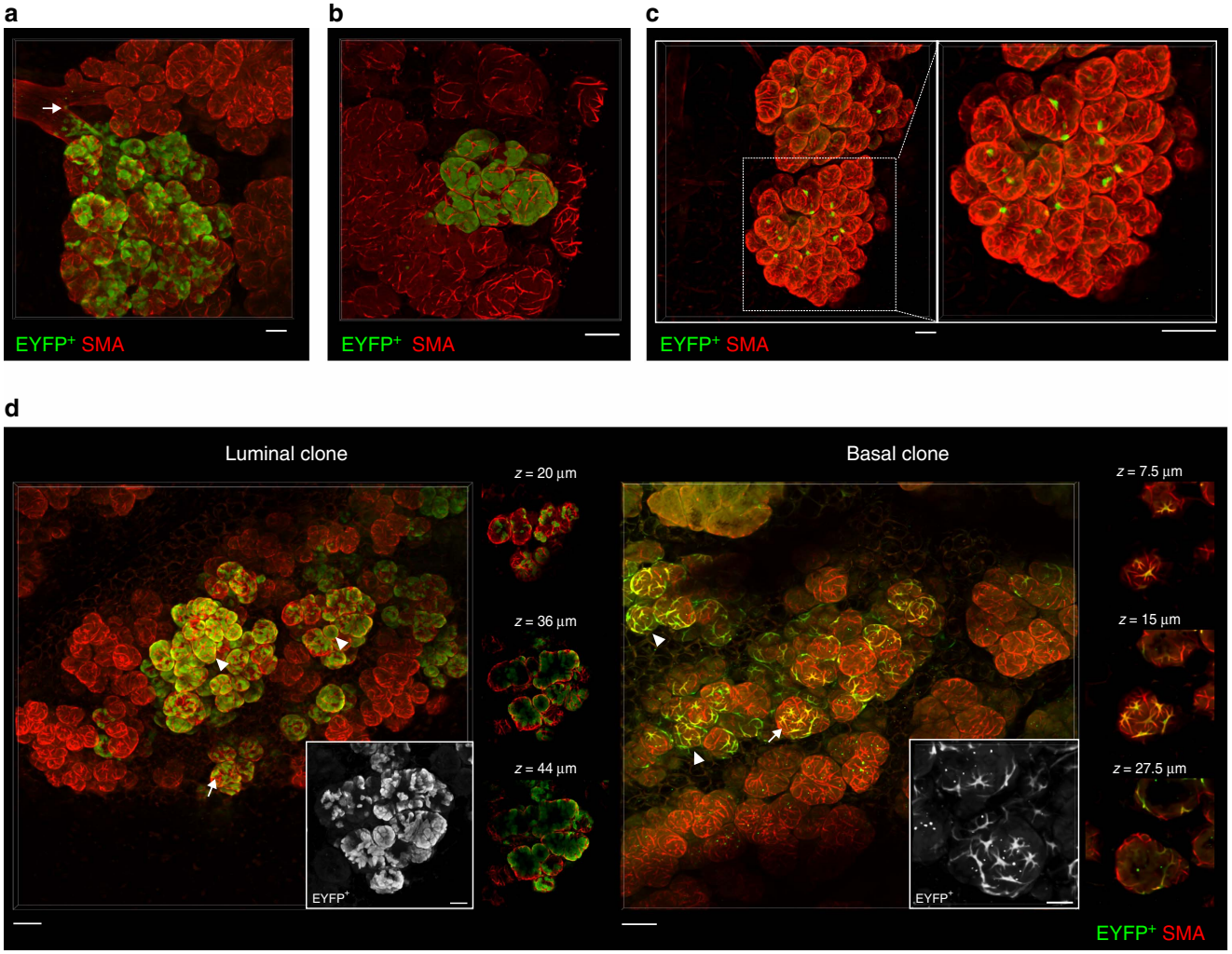
anti-E-cadherin (Cell Signaling, 3,195, 1:50); rabbit anti-PR (DAKO, A0098, 1:50); and chicken anti-GFP (Abcam, ab13970, 1:2,000). The following Alexa Fluor-conjugated secondary antibodies were purchased from Life Technologies and used 1:500: goat anti-mouse 488 (A11001); goat anti-mouse 647 (A21237); goat anti-rat Cy3 (A10522); goat anti-rat 488 (A11006); goat anti-rabbit 488 (A11008); goat anti-rabbit 647 (A21245); goat anti-chicken 488 (A11039); and goat anti-chicken 568 (A11041).

Optical tissue clearing and wholemount immunostaining. Mammary tissue was dissected and cut into large pieces ($\sim 15 \times 15 \times 2$ mm) for immunostaining and clearing. CUBIC-based tissue clearing was performed[13], with minor modifications, for visualization of clones from $R26^{[CA]30SYNbglA}$ mice. CUBIC Reagent 1 was prepared as a mixture of urea (25% w/w), *N*,*N*,*N*′,*N*′-tetrakis(2-hydroxypropyl) ethylenediamine (25% w/w) and Triton X-100 (15% w/w) in distilled water. CUBIC Reagent 2 was prepared using sucrose (44% w/w), urea (22% w/w), 2,2′, 2″-nitrilotriethanol (9% w/w) and Triton X-100 (0.1% w/w) in distilled water. Tissues were immersed in CUBIC Reagent 1 at 37 °C for 3 days. Mammary glands were counterstained for 1.5 h in methyl green (0.5%), washed and de-stained in acid alcohol. Tissues were immersed in CUBIC Reagent 2 at 37 °C for 1–2 days until transparent and imaged using a Leica MZ75 dissecting microscope.

SeeDB-based tissue clearing[12] was combined with wholemount immunolabelling for visualization of fluorescent clones from $R26^{[CA]30EYFP}$ and *R26-Confetti;R26-CreERT2* mice. Mammary tissue was blocked and permeabilized overnight at 4 °C in PBS with Triton X-100 (1%) and bovine serum albumin (10%). Primary antibodies were diluted in blocking buffer and incubated at 4 °C for 4 days with gentle rocking. Tissue was washed and incubated with secondary antibodies for 2 days at 4 °C before further washing and incubation with DAPI (10 µM) for 1–2 h. Samples were serially incubated for 8–16 h at room temperature in 2–3 ml of 20, 40, 60 and 80% (w/v) fructose in distilled water, and subsequently transferred to 100% (w/v) fructose solution (24 h) and 115% (w/v) fructose solution for 24 h or until imaged. All fructose solutions contained α-thioglycerol (0.5%) to inhibit the Maillard reaction[12] and incubations were performed with gentle agitation.

Detection of β-glucosidase expression. For detection of modified β-glucosidase expression[36,37], mammary glands were excised and fixed for 4 h at room temperature in 10% NBF. Endogenous β-glucosidase activity was heat inactivated for 15 min at 65 °C in phosphate-buffered saline (PBS). Wholemount mammary glands were incubated for 48 h at 50 °C in a solution containing 1 part Solution A (5-bromo-6-chloro-3-indolyl-β-D-glucopyranoside (1%) in dimethyl sulfoxide) and 25 parts Solution B (magnesium chloride (0.02% w/v), potassium ferricyanide (0.096% w/v) and potassium ferrocyanide (0.13% w/v) in PBS), with substrate replenishment after 24 h. Mammary glands were post-fixed in 10% NBF overnight at 4 °C. Tissue clearing was performed using the CUBIC clearing protocol. For all animals, ~5 cm of the small intestine distal from the stomach was excised and used as a reaction control for the detection of β-glucosidase expression.

Histology and two-dimensional immunostaining. For histological analysis of tissue from $R26^{[CA]30EYFP}$ mice, SeeDB-based optical tissue clearing was reversed by overnight incubation in PBS at 4 °C. Standard protocols for paraffin processing and embedding using alcohol and xylene were used. Paraffin-embedded sections (6 µm) were de-waxed in xylene, and antigen retrieval was performed by boiling in tri-sodium citrate buffer (10 mM, pH 6.0), for 11 min. Sections were blocked in goat serum (5%) in PBS supplemented with 0.05% Triton X-100 PBS for 1 h at room temperature. Sections were incubated with primary antibodies overnight at 4 °C. Primary antibodies used were as follows: rat anti-K8 (1:200); chicken



e

R26$^{(CA)30EYFP}$				
Total no.	10 mice (80 mammary glands)			
MaSC clone ID	YL.1	YL.2	YL.3	YL.4
Linear length (mm)	>2	>1	>1	>2
No. alveoli	>169	>68	>69	>100
Composition	Lu	Lu	Lu	Ba
Shown in (Fig ref.)	6, S13	S13	S13	6, S13

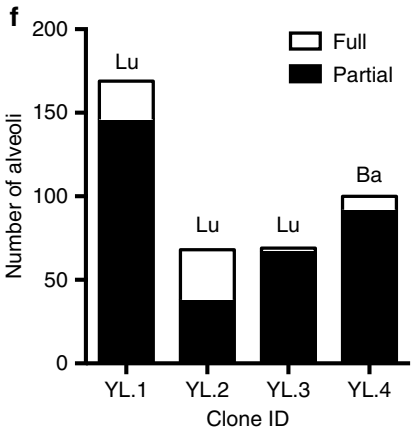


Figure 6 | Contribution of a single MaSC/progenitor to alveologenesis. (**a**) An image showing the uneven contribution of a single labelled ductal cell to different lobuloalveolar structures; arrow indicates the presumptive EYFP$^+$ cell of origin at the ductal branch point. Scale bar, 50 µm. (**b**) A rare instance of progeny from a single luminal EYFP$^+$ cell contributing almost entirely to the luminal lineage of 2–4 alveoli within a single lobule. Scale bar, 50 µm. (**c**) An example of a single labelled EYFP$^+$ luminal cell that contributed one EYFP$^+$ luminal daughter cell to multiple alveoli in a lobule. Scale bar, 100 µm. (**d**) 3D images revealing the extensive contribution of a single luminal (left, clone YL.1) and basal (right, clone YL.4) EYFP$^+$ cell to the lobuloalveolar network in independent lactating mammary glands. Labelled alveoli were mostly populated by both lineage-restricted EYFP$^+$ and unlabelled cells (arrow), with occasional alveoli observed that were fully populated by EYFP$^+$ cells of a single lineage (arrowhead). Scale bar, 100 µm (overview) and 40 µm (inset). (**e**) A summary of the four large clonally marked regions observed from the analysis of 10 *R26*$^{[CA]30EYFP}$ mice during lactation. (**f**) The number of alveoli that were fully populated by EYFP$^+$ cells of a single lineage (full) or populated by both EYFP$^+$ and unlabelled cells of a single lineage (partial). Lu, luminal; Ba, basal.

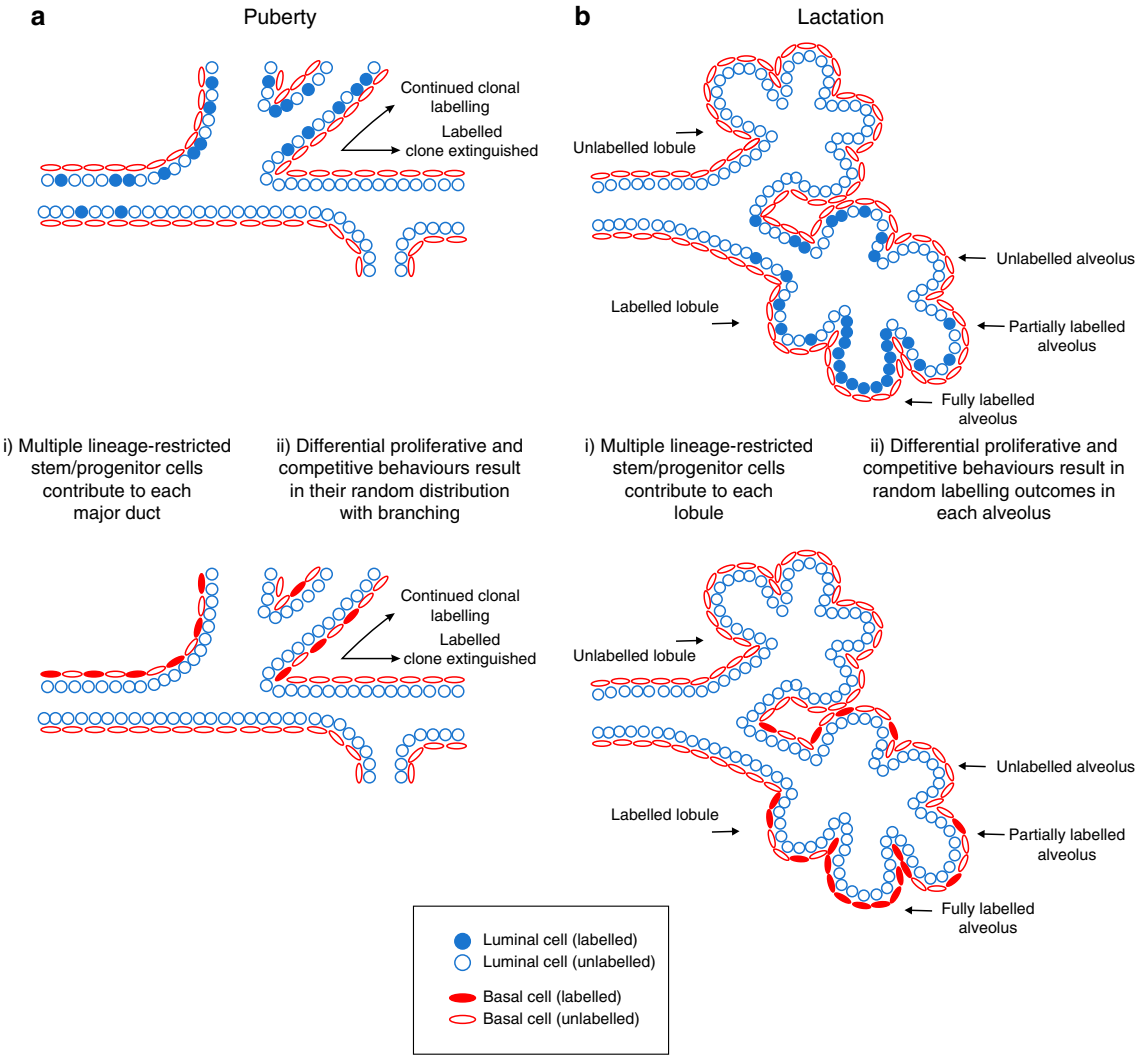


Figure 7 | Schematic model. Schematic model of the most common labelling pattern arising from the genetic labelling of a single lineage-restricted MaSC/progenitor to (**a**) ductal morphogenesis and (**b**) alveologenesis, identified in *R26*$^{[CA]30}$ mice and confirmed using the *R26-Confetti;R26-CreERT2* model (puberty).

anti-GFP (1:2,000); and rabbit anti-SMA (1:200). Secondary antibodies were diluted 1:500. Nuclei were counterstained with DAPI (5 μM).

Confocal microscopy and image analysis. Images of wholemount mammary glands were acquired using a Leica TCS SP8 inverted confocal microscope with 10 × /0.4 or 20 × /0.75 HC PL APO objective lenses. Laser power, line averaging and step increment were adjusted manually to give optimal fluorescence intensity for each fluorophore with minimal photobleaching. Imaging depths were recorded from the top of the epithelial structure being imaged. However, an actual imaging depth of ~350 μm through the native fat pad was typically required before reaching the mammary ductal tree. Thus, imaging depths were routinely 350–500 μm through the tissue. An XZ projection illustrating the lower axial resolution versus the lateral resolution is shown in Supplementary Fig. 14a. The reduced resolution had little effect on the 3D image analyses, with EYFP$^+$ cells able to be discriminated on the anterior and posterior surfaces of most ducts. CFP-expressing clones were rarely observed in *R26-Confetti;R26-CreERT2* mice, possibly related to the brightness of the fluorescent protein, poor penetration of short-wavelength light through the lipid-rich mammary fat pad, the fine membranous localization of the CFP reporter protein, as well as the available laser lines on the confocal microscope. Image reconstructions were generated using Imaris image management software (v8.0, Bitplane). Denoizing of 3D image sequences was performed in MATLAB[14].

Analyses of 3D image stacks, selected on the basis of their resolution and compatibility with 3D image analysis, aimed to identify ducts within the intact mammary stroma and to subsequently recognize all ductal EYFP$^+$ cells. Ductal EYFP$^+$ cells were classified as luminal or basal based on the proportion of K8 versus SMA fluorescence signal. For a volumetric analysis, the volume ratio of EYFP$^+$ cells within each duct was computed with respect to the entire ductal (cellular) volume, and the intensity of K8 in EYFP$^+$ cells was also compared with the overall K8 intensity level within the duct. For computational efficiency, a multi-resolution transform[38] was used for the K8 channel; a coarse scale was used to detect the duct and the full detail scale was used to identify voxels significantly different from the background. The coarse scale was segmented with a robust threshold, obtained as the median of the intensity values in the transformed stack plus three times the median absolute deviation of these values. The up-sampled structure represents an approximation of the duct (Supplementary Fig. 14b) and the sum of all its voxels was a measure of the volume of the duct. Within the duct, significant voxels (excluding intercellular spaces and nuclei) were detected plane-wise from the fine detail levels of the wavelet coefficients of the two-dimensional wavelet transformed image by applying a false discovery rate-based thresholding[39]. Independently, EYFP$^+$ cells were identified after a difference of Gaussian filtering suited to the noise level of the image; as the filter was applied on the full-resolution 3D stack, a recursive filter implementation in CImg (Deriche, CImg) was used for time efficiency. The threshold was computed as above, as the robustly estimated 99% quantile of the Gaussian distribution of filtered intensity values: the median plus three times the median absolute deviation of these intensities (Supplementary Fig. 14c). Subsequently, only EYFP$^+$ cells inside the detected duct were taken into account. These cells were classified as luminal or basal based on the comparison of intensity values in the K8 and SMA channels of the voxels belonging to each segmented cell; for a chosen threshold, the voxels exceeding this threshold in the K8 and SMA channels, respectively, were counted. If the number of K8 voxels exceeded the SMA, the cell was classified as luminal, otherwise it was classified as basal (Supplementary Fig. 14d). Note that a perfect exclusion of one colour cannot be expected due to the resolution limits of the optical system. To provide robustness with respect to

the choice of the intensity threshold, the classification was performed using a multi-threshold approach, with levels 100, 300, 500, up to 1,500 and the majority vote for all thresholds gave the final classification of the cell. Finally, a Kolmogorov–Smirnov test was performed to determine if the significant voxels of the duct in K8 channel were differentially distributed compared with the significant voxels inside the segmented EYFP$^+$ cell.

Some particularities of the basal EYFP$^+$ co-localization images (related to clone YP.3) such as the poorer signal-to-noise ratio of these images (due to the depth of this clone within the mammary fat pad) and the elongated shape of the EYFP$^+$ cells, made a few modifications of the described analysis necessary. To improve the quality of the images, a denoizing step was applied followed by a fast deblurring step (Dr Jerome Boulanger, Medical Research Council—Laboratory of Molecular Biology, private communication). The segmentations are performed in 3D for all channels, however, to separate elongated and overlapping EYFP$^+$ cells, a seeded watershed was used (the seeds are thresholded distance images of the inverted segmented EYFP$^+$ image, where the threshold is manually selected). The classification of the EYFP$^+$ cells is performed as before: if the number of K8 voxels exceeded the SMA, the cell was classified as luminal, otherwise it was classified as basal. Using this analysis, two cells were excluded from the classification due to their localization in regions where the SMA signal was undetectable and thus the double/nested tubular structure could not be observed.

Quantification of PR$^+$ and K8^{hi} cells (Supplementary Fig. 3b) was performed on maximum intensity projections of 3D image stacks using the Cell Counter plugin in Image J (v1.50a, National Institutes of Health). Maximum intensity projections of PR and K8 channels were scored independently. At least 200 cells were counted per image, with six images analysed from three independent mice (total cells counted: 1,831). Manual counting of EYFP$^+$ cells (Fig. 3d and Supplementary Fig. 10b) was performed on 10 histological sections cut $>25\,\mu m$ apart, and spanning $\sim 300\,\mu m$. K8 was used to mark the luminal lineage and SMA was used to mark basal cells.

The number of alveoli that were fully or partially populated by EYFP$^+$ cells of a single lineage were manually counted in Image J.

Statistics. The Kolmogorov–Smirnov test (Supplementary Fig. 10a) was performed in MATLAB (R2014a, The MathWorks Inc., Natick, Massachusetts). All values are shown as mean ± s.d.

Data availability. The data supporting the findings of this study are available within the article and its Supplementary Information files. All other relevant source data are available from the authors on request.

References

1. Watson, C. J. & Khaled, W. T. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development* **135,** 995–1003 (2008).
2. Sreekumar, A., Roarty, K. & Rosen, J. M. The mammary stem cell hierarchy: a looking glass into heterogeneous breast cancer landscapes. *Endocr. Relat. Cancer* **22,** T161–T176 (2015).
3. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439,** 84–88 (2006).
4. Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439,** 993–997 (2006).
5. Sleeman, K. E., Kendrick, H., Ashworth, A., Isacke, C. M. & Smalley, M. J. CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast Cancer Res.* **8,** R7 (2006).
6. Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479,** 189–193 (2011).
7. Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. *In situ* identification of bipotent stem cells in the mammary gland. *Nature* **506,** 322–327 (2014).
8. van Amerongen, R. *et al.* Developmental stage and time dictate the fate of Wnt/β-Catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* **11,** 387–400 (2012).
9. Wang, D. *et al.* Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517,** 81–84 (2015).
10. de Visser, K. E. *et al.* Developmental stage-specific contribution of LGR5(+) cells to basal and luminal epithelial lineages in the postnatal mammary gland. *J. Pathol.* **228,** 300–309 (2012).
11. Richardson, D. S. & Lichtman, J. W. Clarifying tissue clearing. *Cell* **162,** 246–257 (2015).
12. Ke, M.-T., Fujimoto, S. & Imai, T. SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nat. Neurosci.* **16,** 1154–1161 (2013).
13. Susaki, E. A. *et al.* Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* **157,** 726–739 (2014).
14. Boulanger, J. *et al.* Patch-based nonlocal functional for denoising fluorescence microscopy image sequences. *IEEE Trans. Med. Imaging* **29,** 442–454 (2010).
15. Sleeman, K. E. *et al.* Dissociation of estrogen receptor expression and *in vivo* stem cell activity in the mammary gland. *J. Cell Biol.* **176,** 19–26 (2007).
16. Koren, S. *et al.* PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours. *Nature* **525,** 114–118 (2015).
17. Kozar, S. *et al.* Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell* **13,** 626–633 (2013).
18. Williams, J. M. & Daniel, C. W. Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. *Dev. Biol.* **97,** 274–290 (1983).
19. Strickland, P., Shin, G. C., Plump, A., Tessier-Lavigne, M. & Hinck, L. Slit2 and netrin 1 act synergistically as adhesive cues to generate tubular bi-layers during ductal morphogenesis. *Development* **133,** 823–832 (2006).
20. Paine, I. *et al.* A geometrically-constrained mathematical model of mammary gland ductal elongation reveals novel cellular dynamics within the terminal end bud. *PLoS Comput. Biol.* **12,** e1004839 (2016).
21. Zeps, N., Bentel, J. M., Papadimitriou, J. M. & Dawkins, H. J. Murine progesterone receptor expression in proliferating mammary epithelial cells during normal pubertal development and adult estrous cycle. Association with eralpha and erbeta status. *J. Histochem. Cytochem.* **47,** 1323–1330 (1999).
22. Giraddi, R. R. *et al.* Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. *Nat. Commun.* **6,** 8487 (2015).
23. Spike, B. T. *et al.* A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell Stem Cell* **10,** 183–197 (2012).
24. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450,** 56–62 (2007).
25. Ventura, A. *et al.* Restoration of p53 function leads to tumour regression *in vivo*. *Nature* **445,** 661–665 (2007).
26. Wuidart, A. *et al.* Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes Dev.* **30,** 1261–1277 (2016).
27. Johnston, L. A. Competitive interactions between cells: death, growth, and geography. *Science* **324,** 1679–1682 (2009).
28. Klein, A. M. & Simons, B. D. Universal patterns of stem cell fate in cycling adult tissues. *Development* **138,** 3103–3111 (2011).
29. DeOme, K. B., Faulkin, L. J., Bern, H. A. & Blair, P. B. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer Res.* **19,** 515–520 (1959).
30. Boras-Granic, K., Dann, P. & Wysolmerski, J. J. Embryonic cells contribute directly to the quiescent stem cell population in the adult mouse mammary gland. *Breast Cancer Res.* **16,** 487 (2014).
31. Sun, P., Yuan, Y., Li, A., Li, B. & Dai, X. Cytokeratin expression during mouse embryonic and early postnatal mammary gland development. *Histochem. Cell Biol.* **133,** 213–221 (2010).
32. Amoyel, M., Simons, B. D. & Bach, E. A. Neutral competition of stem cells is skewed by proliferative changes downstream of Hh and Hpo. *EMBO J.* **33,** 2295–2313 (2014).
33. Morrissey, E. R. & Vermeulen, L. Stem cell competition: how speeding mutants beat the rest. *EMBO J.* **33,** 2277–2278 (2014).
34. dos Santos, C. O. *et al.* Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proc. Natl Acad. Sci. USA* **110,** 7123–7130 (2013).
35. Kordon, E. C. & Smith, G. H. An entire functional mammary gland may comprise the progeny from a single cell. *Development* **125,** 1921–1930 (1998).
36. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330,** 822–825 (2010).
37. McCutcheon, S. C. *et al.* Characterization of a heat resistant beta-glucosidase as a new reporter in cells and mice. *BMC Biol.* **8,** 89 (2010).
38. Adelson, E. H., Simoncelli, E. P. & Freeman, W. T. *Pyramids and Multiscale Representations. Representation of Vision: Trends and Tacit Assumptions in Vision Research* (Cambridge Univ. Press, 1991).
39. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57,** 289–300 (1995).

Acknowledgements

We thank the Department of Pathology Biological Services Unit for help with animal work, H. Skelton for help with histology and Peter Humphreys and Dr Martin Lenz for imaging support. This work was supported by a grant from the Medical Research Council programme grant no. MR/J001023/1 (B.L-L. and C.J.W). F.M.D. was funded by a National Health and Medical Research Council CJ Martin Biomedical Fellowship (GNT1071074). O.B.H. was funded by a Wellcome Trust PhD studentship (105377/Z/14/Z).

Author contributions

B.L-L., F.M.D. and O.B.H. performed all experiments; B.L.-L., F.M.D., O.B.H., S.K., D.J.W. and C.J.W. conceived and designed the experiments; L.M. designed and performed the 3D image analysis; B.L.-L., F.M.D and C.J.W. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Davis, F. M. *et al.* Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat. Commun.* **7,** 13053 doi: 10.1038/ncomms13053 (2016).

Lloyd-Lewis et al. Breast Cancer Research (2016) 18:127
DOI 10.1186/s13058-016-0754-9

Breast Cancer Research

RESEARCH ARTICLE Open Access

Imaging the mammary gland and mammary tumours in 3D: optical tissue clearing and immunofluorescence methods

Bethan Lloyd-Lewis[1*†], Felicity M. Davis[1,2*†], Olivia B. Harris[1,3], Jessica R. Hitchcock[1], Filipe C. Lourenco[4], Mathias Pasche[5] and Christine J. Watson[1,3*]

Abstract

Background: High-resolution 3D imaging of intact tissue facilitates cellular and subcellular analyses of complex structures within their native environment. However, difficulties associated with immunolabelling and imaging fluorescent proteins deep within whole organs have restricted their applications to thin sections or processed tissue preparations, precluding comprehensive and rapid 3D visualisation. Several tissue clearing methods have been established to circumvent issues associated with depth of imaging in opaque specimens. The application of these techniques to study the elaborate architecture of the mouse mammary gland has yet to be investigated.

Methods: Multiple tissue clearing methods were applied to intact virgin and lactating mammary glands, namely 3D imaging of solvent-cleared organs, see deep brain (seeDB), clear unobstructed brain imaging cocktails (CUBIC) and passive clarity technique. Using confocal, two-photon and light sheet microscopy, their compatibility with whole-mount immunofluorescent labelling and 3D imaging of mammary tissue was examined. In addition, their suitability for the analysis of mouse mammary tumours was also assessed.

Results: Varying degrees of optical transparency, tissue preservation and fluorescent signal conservation were observed between the different clearing methods. SeeDB and CUBIC protocols were considered superior for volumetric fluorescence imaging and whole-mount histochemical staining, respectively. Techniques were compatible with 3D imaging on a variety of platforms, enabling visualisation of mammary ductal and lobulo-alveolar structures at vastly improved depths in cleared tissue.

Conclusions: The utility of whole-organ tissue clearing protocols was assessed in the mouse mammary gland. Most methods utilised affordable and widely available reagents, and were compatible with standard confocal microscopy. These techniques enable high-resolution, 3D imaging and phenotyping of mammary cells and tumours *in situ*, and will significantly enhance our understanding of both normal and pathological mammary gland development.

Keywords: Mammary gland, Lactation, Breast cancer, Tissue clearing, 3D imaging, Fluorescence microscopy, Light sheet fluorescence microscopy, Two-photon microscopy

* Correspondence: bl349@cam.ac.uk; fmd32@cam.ac.uk; cjw53@cam.ac.uk
†Equal contributors
[1]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK
Full list of author information is available at the end of the article

Background

The mammary gland is composed of a branching epithelial ductal network deeply embedded within a vascularised stromal matrix made up of adipocytes, fibroblasts and immune cells [1]. Due to its capacity for rapid growth and regeneration, the mouse mammary gland is a powerful model in which to study a range of developmental processes associated with tissue morphogenesis and remodelling, and provides important insights into the perturbations that give rise to breast cancer [1]. However, visualisation of the complex cellular networks within the intact mammary epithelium is greatly impeded by the lipid-rich and optically opaque nature of this organ. As a result, immunolabelling and fluorescence imaging of mammary tissue has traditionally been performed using thin tissue sections with assumptions about the architectural context and uniformity of a particular 2D plane. Whilst 3D imaging has recently been used to investigate mammary stem cell dynamics [2–4] and binucleated cells in lactational alveoli [5], these studies have relied on tissue microdissection [2, 5] or enzymatic digestion [3]. Consequently, visualisation of the mammary epithelial tree at single-cell resolution within its native stroma remains a fundamental challenge in mammary gland research.

The utility of rendering tissue optically transparent has been appreciated for over a century [6]. However, recent advances in fluorescence microscopy have heralded the development of numerous whole-organ tissue clearing methods aimed at improving optical access and depth of imaging in intact specimens (reviewed in [7]). These methods are primarily based on mitigating light scattering caused by heterogeneous cellular components with different refractive indices (RIs). Techniques broadly rely either on organic solvent-based or hydrophilic reagent-based clearing solutions to homogenise RIs within tissue, and may also include prior hydrogel embedding to stabilise cellular structures. Whilst many protocols were originally optimised for the central nervous system and whole embryos, recent refinements in tissue clearing techniques have facilitated exceptional optical access to many other mammalian tissues. However, the application of these techniques to the mammary gland is yet to be explored.

Here, we describe the application of four leading tissue clearing protocols, namely three-dimensional imaging of solvent-cleared organs (3DISCO) [8], see deep brain (SeeDB) [9], clear unobstructed brain imaging cocktails (CUBIC) [10] and passive clarity technique (PACT) [11], to virgin and lactating mammary glands. Whilst the underlying principles for achieving tissue transparency are fundamentally different in each of these methods, the majority utilise simple and affordable reagents, and can be completed within two weeks.

3DISCO [8] is based on earlier clearing methods that use high-index organic solvents such as benzyl alcohol benzyl benzoate (BABB) [12, 13], and can be combined with optimised whole-mount immunolabelling procedures (termed iDISCO) [14]. To date, 3DISCO remains the only method previously applied to mammary tissue, albeit superficially [8]. PACT [11] relies on hydrogel embedding to stabilise cellular structures prior to tissue delipidation using ionic detergents. Due to the need for custom electrophoresis equipment, CLARITY [15] is difficult to implement and can lead to heat-induced tissue damage and epitope loss [11]. The PACT protocol circumvents these issues, relying instead on passive diffusion. Furthermore, PACT utilises more economical RI matching solutions (RIMS and sRIMS), an additional benefit over CLARITY. CUBIC [10] is a urea-based clearing reagent that includes aminoalcohols and detergents to remove lipids and homogenise RIs within tissue. In addition, CUBIC reagents decolourise blood by eluting the heme chromophore, further enhancing optical transparency by minimising light absorption [16]. Finally, SeeDB [9] is a water-based optical clearing agent that utilises saturated solutions of fructose and alpha-thioglycerol (which limits autofluorescence) for RI matching.

We compared these four protocols for optical transparency and tissue preservation in the intact mammary gland, in addition to their compatibility with immunofluorescent labelling and 3D imaging of mammary epithelial cells. Their suitability for the analysis of mammary tumours was also investigated. Using standard confocal and advanced imaging techniques, ductal and lobulo-alveolar structures could be readily visualised in cleared tissue, with varying degrees of fluorescent signal preservation between the different methods. Overall, our results placed SeeDB and CUBIC as methods of choice for high-resolution fluorescence imaging and whole-mount histochemical staining of mammary glands. The ability to visualise the mammary epithelial tree at single cell resolution within its native stroma will provide invaluable insight into mammary gland development and tumourigenesis.

Methods

Reagents and antibodies

The following reagents were purchased from Sigma Aldrich: neutral buffered formalin (NBF), dimethyl sulfoxide (DMSO), tetrahydrofuran (THF), dichloromethane (DCM), dibenzyl ether (DBE), benzyl alcohol, benzyl benzoate, urea, N,N,N′,N′-tetrakis(2-hydroxypropyl)ethylenediamine, 2,2′,2″-nitrilotriethanol, fructose, α-thioglycerol, D-sorbitol and 4′,6-diamidino-2-phenylindole (DAPI) dilactate. RapiClear-CLARITY Specific (RC-CS) Solution and Mounting Medium and iSpacer imaging chambers were purchased from the SunJin Lab: 2,2′-Azobis[2-(2-imidazolin-2-yl)propane]

dihydrochloride was purchased from Wako Pure Chemical Industries. Imaging dishes were purchased from Ibidi (81158). Acrylamide (40%) was purchased from Bio-Rad Laboratories. Sodium dodecyl sulphate (SDS) was purchased from Melford Laboratories. Triton-X100 was purchased from VWR International. The following primary antibodies were used for immunostaining: rabbit anti-α-smooth muscle actin (SMA) (Abcam, ab5694; 1:200-1:300 for 2D and 3D studies), rabbit anti-keratin 5 (BioLegend, 905501; 1:100 (3D)), rat anti-cytokeratin 8 (Developmental Studies Hybridoma Bank, TROMA-I; 1:50 (3D) or 1:150-200 (2D)), rabbit anti-E-cadherin (Cell Signaling, 3195; 1:50 (3D) or 1:200 (2D)), mouse anti-E-cadherin (BD Transduction Laboratories, 610182; 1:300 (2D)), rabbit anti-cleaved caspase 3 (Cell Signaling, 9661S; 1:200 (2D)) and rabbit anti-human epidermal growth factor receptor 2 (HER2) (DAKO, A0485; 1:300 (3D) or 1:500 (2D)). The following Alexa Fluor conjugated secondary antibodies were purchased from Life Technologies and diluted 1:500 (2D and 3D studies) in blocking buffer: goat anti-rat Cy3 (A10522), goat anti-rabbit 647 (A21245) and chicken anti-rabbit 647 (A21443). Anti-rabbit horseradish peroxidase (HRP)-conjugated secondary antibody was purchased from DAKO (P0448; 1:500).

Mice

Mice (C57BL/6 and BALB/c) were housed in individually ventilated cages under a 12:12 h light-dark cycle, with water and food available *ad libitum*. Mice were euthanized by dislocation of the neck or terminal anaesthesia. All tissue from virgin mice was harvested during puberty (5–8 weeks). For studies during lactation, mice were mated with studs, allowed to litter and tissue was harvested between lactation days 2 to 10. Mammary glands (excluding the cervical (first) pair) were excised and immediately spread and fixed on card (Tetra Pak) in 10% NBF for 9 h at room temperature, unless otherwise specified. A 9-h fix provided optimal staining for all antibodies used in this study; however, SMA and K8 also performed well with overnight fixation (4 °C). Fixed tissue was stored at 4 °C in phosphate-buffered saline (PBS) containing sodium azide (0.05% (w/v)) for up to 8 weeks.

Syngeneic mammary tumours were established by orthotopically implanting 5×10^3 TUBO cells [17], into the abdominal (fourth) mammary gland. This cloned cell line was established from a mammary carcinoma that spontaneously arose in a BALB-neuT mouse and therefore carries the Her-2/neu oncogene driven by the MMTV promoter. Mice were monitored regularly and tumours were harvested before exceeding humane endpoints (approx. 4–5 weeks).

3DISCO-based clearing and immunohistochemical analysis (IHC)

3DISCO was performed as previously described [8]. Solvent immersion times were adjusted for mammary tissue pieces (approx. 10 × 10 × 1 mm) as follows: 50% (v/v) THF in H_2O (40 min), 70% (v/v) THF in H_2O (40 minutes), 80% (v/v) THF in H_2O (40 minutes), 100% (v/v) THF (3 × 40 minutes), and DCM (15 minutes). All solvent immersion steps were performed in glass vials. Although DBE is reported to be a superior optical clearing agent vs. BABB [8], this solvent requires specialised imaging chambers and solvent-resistant adhesive (e.g. dental cement), and can severely damage objectives if the chamber fails. Thus, we used BABB as the final clearing agent for transmission and confocal imaging in this study. BABB was prepared as a mixture of benzyl alcohol and benzyl benzoate (1:2). Immunostaining was performed as per the iDISCO protocol [14], with a methanol pretreatment. Following the iDISCO immunolabelling protocol, samples were washed in PBS and incubated with DAPI (10 μM) for 2–3 h, cleared using the 3DISCO protocol and imaged the same day.

PACT-based clearing and IHC

The A4P0 hydrogel formulation was selected for PACT-based clearing of the mouse mammary gland [11]. A4P0 was prepared to contain acrylamide (4% (v/v)), 2,2′-Azobis[2-(2-imidazolin-2-yl)propane] dihydrochloride (0.25% (w/v)) in PBS. PACT-clearing solution consisted of SDS (8% (w/v)) in distilled water, pH 7.5. Mammary tissue pieces (approx. 10 × 10 × 1 mm) were incubated in A4P0 hydrogel monomer for 4 days at 4 °C and heated to 37 °C in a water bath for 4–6 h. Excess gel was carefully removed from the tissue and samples were immersed in PACT clearing solution for 24 h at room temperature. Samples were immersed in fresh clearing solution, incubated at 37 °C for 4 days (with replenishment every second day), and finally washed with PBS containing triton-X100 (0.1% (w/v)) for 24 h. For immunostaining, samples were blocked in PBS containing triton-X100 (0.5% (w/v)) with goat serum (10% (v/v)) overnight at 4 °C. Primary antibodies were diluted in blocking buffer at 4 °C for 4 days with agitation, tissue was washed (3 × 1 h) in PBS and incubated with Alexa Fluor conjugated secondary antibodies for 2 days. Samples were washed in PBS and incubated with DAPI (10 μM) for 2–3 h. PACT-sRIMS samples were incubated in sRIMS for 4 days or until imaging. sRIMS was prepared by combining sorbitol (70% (w/v)) in 0.02 M phosphate buffer [11, 18]. PACT-RC samples were incubated in Rapiclear CS for 4 h and mounted between two coverslips using RC-CS Mounting Medium and iSpacers for image acquisition and long-term storage.

CUBIC-based clearing and IHC

CUBIC-based tissue clearing was performed as previously described [10], with minor modifications. CUBIC Reagent 1 was prepared as a mixture of urea (25% (w/w)), N,N,N',N'-tetrakis(2-hydroxypropyl)ethylenediamine (25% (w/w)), triton-X100 (15% (w/w)) in distilled water. CUBIC Reagent 2 was prepared using sucrose (44% (w/w)), urea (22% (w/w)), 2,2′,2″-nitrilotriethanol (9% (w/w)), triton-X100 (0.1% (w/w)) in distilled water. CUBIC Reagent 1A was prepared using urea (10% (w/w))), N,N,N',N'-tetrakis(2-hydroxypropyl)ethylenediamine (5% (w/w)), triton-X100 (10% (w/w)) and NaCl (25 mM) in distilled water (unpublished, protocol available at http://cubic.riken.jp/). Tissue pieces (approx. 10 × 10 × 1 mm) were immersed in CUBIC Reagent 1 or 1A at 37 °C for 2–3 days, depending on the size of the tissue piece. For immunostaining samples were washed and subsequently blocked in PBS containing triton-X100 (0.5% (w/v)) and goat serum (10% (v/v)) overnight at 4 °C. Primary antibodies were diluted in blocking buffer at 4 °C for 4 days with gentle rocking. Tissue was washed (3 × 1 h) and incubated with Alexa Fluor conjugated secondary antibodies for 2 days, washed in PBS and incubated with DAPI (10 μM) for 2–3 h. Samples were transferred to CUBIC Reagent 2 at 37 °C for at least 1 day for refractive index matching. Samples were immersed in CUBIC Reagent 2 for imaging and were imaged within 1 week. Diffuse, non-specific fluorescence was observed using the CUBIC protocol in the absence of the primary antibodies (Additional file 1: Figure S1, top panel).

Whole-mount histochemical and IHC analysis combined with CUBIC-based tissue clearing

Excised and fixed mammary glands were immersed in reagent 1 for 2–3 days. Glands were removed and stained with methyl green (0.5% (w/v)), Harris haematoxylin (10%) or carmine for 1.5–2 h at room temperature with gentle agitation. After staining, tissues were rinsed twice in tap water and once in distilled water before de-staining with acid alcohol (50% ethanol with hydrochloric acid (25 mM)) for 20 minutes and immersion in reagent 2. For detection of β-glucosidase expression (magenta histochemical stain) [19], mammary glands were excised and fixed for 4 h at room temperature. Endogenous β-glucosidase activity was heat inactivated at 65 °C for 15 minutes in PBS. Whole mammary glands were incubated for 48 h at 50 °C in a solution containing 1 part Solution A (5-Bromo-6-chloro-3-indolyl- β-D-glucopyranoside (1% (w/v)) in DMSO) and 25 parts solution B (magnesium chloride (0.02% (w/v)), potassium ferricyanide (0.096% (w/v)) and potassium ferrocyanide (0.13% (w/v)) in PBS). Mammary glands were post-fixed in 10% NBF overnight at 4 °C and cleared using the standard CUBIC-clearing protocol.

For whole-mount immunohistochemical analysis, samples were dehydrated by a methanol series and incubated overnight in methanol containing DMSO (20%) and H_2O_2 (3%) to quench endogenous peroxidase activity [20]. Samples were rehydrated by methanol series and blocked and permeabilised in PBS containing BSA (10% (w/v)) and triton-X100 (1% (w/v)). Samples were incubated with rabbit anti-SMA antibody (1:200) for 4 days at 4 °C with gentle agitation. After washing, samples were incubated with anti-rabbit HRP-conjugated secondary antibody (1:500) for a further 2 days, before immersion in reagent 2. Alternatively, quenching, blocking and antibody steps can be performed after immersion in reagent 1 with a similar outcome.

SeeDB-based clearing and IHC

SeeDB-based clearing was performed as previously described [9], with minor modifications. Briefly, mammary tissue pieces (approx. 10 × 10 × 1 mm) were blocked and permeabilised overnight at 4 °C in PBS with triton-X100 (1% (w/v)) and BSA (10% (w/v)). Primary antibodies were diluted in blocking buffer at 4 °C for 4 days with gentle rocking. Tissue was washed (3 × 1 h) and incubated with secondary antibodies for 2 days before further washing in PBS and incubation with DAPI (10 μM) for 2–3 h. Samples were serially incubated for 8–16 h (twice daily changes) in 2–3 mL of 20%, 40%, 60% and 80% (w/v) fructose in distilled water, and subsequently transferred to 100% (w/v) fructose solution (24 h) and 115% (w/v) fructose solution for 24 h or until imaging. All fructose solutions contained α-thioglycerol (0.5% (v/v)) to inhibit the Maillard reaction [9, 21] and incubations were performed with gentle agitation. For optimal performance, samples were imaged within 2 weeks of clearing; however, staining was still observed up to 6 months after clearing. Diffuse, non-specific fluorescence was observed using the SeeDB protocol in the absence of the primary antibodies (Additional file 1: Figure S1, bottom panel).

Two-dimensional IHC on CUBIC-recovered and SeeDB-recovered tissue, and formalin-fixed paraffin-embedded (FFPE) tumour sections

Samples were immersed in PBS for approx. 3 days for passive rehydration and removal of the clearing agent [9]. Standard protocols for paraffin processing and embedding using alcohol and xylene were employed. Paraffin-embedded sections (4–6 μm) were de-waxed in xylene and antigen retrieval was performed by boiling in a pressure cooker in tri-sodium citrate buffer (10 mM, pH 6.0), for 11 minutes [22]. Sections were blocked in goat serum (5% (v/v)) in PBS supplemented with triton-X100 (0.05% (w/v)) for 1 h in a humidified chamber at room temperature. Sections were incubated with primary antibodies overnight at 4 °C. Primary antibodies used were: rat anti-K8, rabbit anti-SMA, rabbit anti-E-

cadherin, mouse anti-E-cadherin and rabbit anti-cleaved caspase 3. Alexa Fluor conjugated secondary antibodies were diluted 1:500. Nuclei were counterstained with DAPI (1–5 μM).

Optical clearing and measurement of sample size changes

Mammary tissue pieces were processed using 3DISCO, PACT-RC, PACT-sRIMS, CUBIC or SeeDB-based tissue clearing protocols, and images were acquired on a dissecting microscope (Leica MZ75) with constant exposure, gain and magnification. For quantification of sample size changes, image thresholding was performed using ImageJ (v1.49p, National Institutes of Health) and the pixel area was measured [9]. Volume changes were calculated as the ratio of the pixel area before and after tissue clearing.

Confocal microscopy

Tissues cleared by 3DISCO, PACT-sRIMS, CUBIC and SeeDB were imaged in their respective RI matching solutions in Ibidi μ-Dishes. PACT-RC-cleared tissues were mounted using iSpacer chambers in RC-CS Mounting Medium. Images were acquired using a Leica TCS SP8 inverted confocal microscope with 10×/0.4 or 20×/0.75 HC PL APO objective lenses. Laser power and gain were adjusted manually to give optimal fluorescence intensity for each fluorophore with minimal photobleaching. Step size and line averaging were kept constant for all main figures (line averaging, 16; step size 1–2), excluding CUBIC and SeeDB depth cueing examples. Imaging depths were recorded from the top of the epithelial structure being imaged (typically 350 μm through the native fat pad for the CUBIC and SeeDB protocols). Image reconstructions were generated using Imaris image management software (v8.0, Bitplane) or ImageJ (v1.50c, National Institutes of Health) [23, 24]. Depth coding was performed using the Temporal Colour Code plugin with the spectrum LUT. De-noising of 3D image sequences was performed in MATLAB (R2014a, The Mathworks Inc.) [25].

Two-photon and light sheet fluorescence microscopy (LSFM)

Two-photon imaging was performed on a LaVision BioTec TriM Scope II with a 25×/1.05 water dipping lens and an insight DeepSee dual-line laser (tuneable 710–1010 nm and fixed 1040 nm lines), with 810 nm wavelength used to excite DAPI and HER2-AF647. Tiled images, having a 20% overlap, were stitched together using the Grid Collection/Stitching plugin in ImageJ [26].

For LSFM, samples were immunostained and cleared according to the CUBIC protocol [10]. After clearing, samples were embedded within an agarose (1% (w/v) in H_2O) tube, prepared by aspirating agarose (37–38 °C) into a pre-warmed 1-mL syringe in which the syringe neck had been cut off. Mammary tissue strips were quickly placed within the agarose tube using forceps and centred by rolling the syringe between the palms. After setting, the plunger was removed and the entire syringe was submerged in CUBIC reagent 2. Samples were imaged in reagent 2 or glycerol in H_2O (34% (w/w)).

The light sheet system was a home-built modified version of the OpenSPIM system [27]. The microscope was built and operated in the T-SPIM layout, whereby illumination happens from two sides simultaneously by overlapping two individual sheets to allow a more even illumination and to reduce artefacts, such as striping. We used two Olympus 5×/0.15 air lenses to generate the light sheet. The higher refractive indices, long working distance (20 mm) and the fact that the lenses were on threaded mounts allowed us to adjust the point of focus accordingly. The imaging light path was equipped with a Nikon 16×/0.8 water dipping lens. We imaged onto an Andor Neo 5.5 (ANDOR) or a Hamamatsu ORCA-Flash4 V2 (Hamamatsu) with 6.5 um pixels. For excitation a home-built laser combiner was used, bundling 405 nm, 488 nm, 561 nm and 640 (Coherent Cube 405 and 640, Coherent Sapphire 488 and 561) into a single-mode fibre. Channels were acquired sequentially and emission was filtered by suitable band-pass or long-pass filters (DAPI: 447/60; AF647: 705/72; both AHF Analysentechnik). The sample was mounted in a 4D (xyzθ) stage (Picard Industries) allowing optimal positioning of the sample in the light sheet. During imaging the sample was moved through the light sheet with a step size of 1.5 μm and the light sheet thickness was adjusted to be ca. 6 μm to warrant an even thickness of the sheet across the entire sample width. Exposure times were between 15 and 150 ms.

Results

Optical transparency in the mammary gland

We evaluated four passive whole-organ tissue clearing protocols for optical clarity and morphology preservation during two distinct phases of mammary gland development, puberty and lactation. The PACT protocol was evaluated using two RI matching solutions: the commercially available aqueous-based solution RapiClear CS® (RI = 1.45) and a more-economical sorbitol-based solution, sRIMS (RI = 1.46) [18]. Of the four methods tested, PACT-based protocols [11] were the most time-intensive and labour-intensive, taking between 10 and 13 days for completion (Fig. 1a). In contrast, the simple, immersion-based tissue clearing protocols CUBIC [10] and SeeDB [9] required only 5 days (Fig. 1a), and resulted in superior optical clarity in both virgin and lactating tissue (Fig. 1b). CUBIC clearing provided the highest degree of transparency in mammary tissue and was also highly effective in decolourising blood vessels (Fig. 1b) [16]. To determine whether these methods altered the structural integrity of mammary

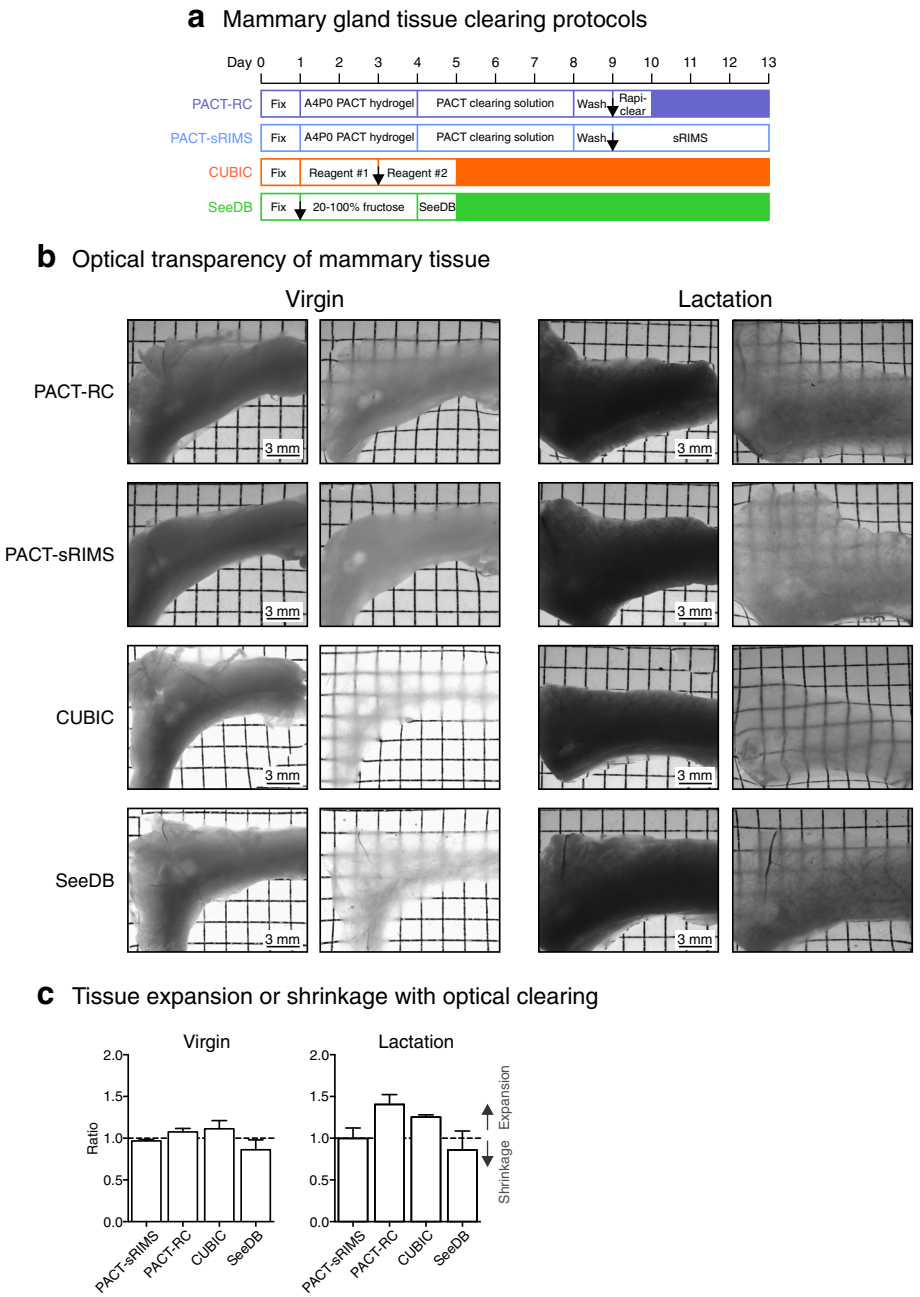


Fig. 1 Optical clearing of mammary tissue. **a** Experimental procedure and timelines for optical clearing of mouse mammary tissue. *Black arrows* show the stage at which (optional) immunostaining may be performed. The experimental timeline can be altered depending on the size and nature of the tissue, and the degree of transparency required. **b** Transmission images of whole abdominal (fourth) mammary glands (virgin and lactating) before and after clearing using the passive clarity technique (*PACT*) with rapiclear (*RC*) or sorbitol refractive index matching solution (*sRIMS*), clear unobstructed brain imaging cocktails (*CUBIC*) or the see deep brain (*SeeDB*) clearing protocols. **c** Volume changes resulting from optical clearing of virgin and lactating mammary tissue. Values are representative of measurements from three tissue pieces from each clearing protocol at each developmental time point. See additional file 18 for a high resolution version of these PDFs and http://rdcu.be/lT3Z for additional high resolution examples of these imaging techniques [38]

tissue, as has been reported for other clearing protocols [12, 28], we measured the sample volume of fixed mammary tissue before and after tissue clearing protocols were applied (Fig. 1c). PACT-RC and CUBIC were associated with a moderate degree of tissue expansion, which was more prominent during lactation. A small reduction in sample volume was observed with SeeDB (Fig. 1c).

We also tested the solvent-based tissue clearing protocol 3DISCO (Additional file 2: Figure S2a), a method previously employed for tissue clearing in the mammary gland [8]. Using 3DISCO, we observed shrinkage of both virgin and lactating mammary tissue samples, which was associated with significant structural deformations to ducts and lactational alveoli (Additional file 2: Figure S2b-d). Although this protocol was extremely rapid to implement (Additional file 2: Figure S2a), achieved a high degree of optical clarity (Additional file 2: Figure S2b) and may be useful in other organ systems [8], its application to the mammary gland, which was left brittle and damaged from the clearing process, is extremely limited. Additionally, the solvents used in this protocol pose significant laboratory safety risks and require specialised imaging chambers [8]. For these reasons we did not pursue 3DISCO further for optical clearing of mammary tissue.

PACT-based tissue clearing and 3D imaging of the unsectioned mouse mammary gland

We assessed PACT-based clearing approaches for the visualisation of mammary epithelial cells in virgin and lactating mammary glands *in situ*. When combined with whole-mount immunostaining, these protocols required 17 (PACT-RC) to 20 (PACT-sRIMS) days preparation prior to imaging (Additional file 3: Figure S3a and Fig. 2a) [11]. Samples prepared using the PACT-RC protocol are not stable in the RI matching solution, and thus require mounting in a specialised mounting medium (SunJin Labs). Whilst this approach is conducive to long-term sample storage, the limited working distances of standard confocal microscope objectives makes imaging of PACT-RC-mounted samples problematic, as samples cannot be re-orientated against the coverglass for optimal sample illumination (Additional file 3: Figure S3b-c). Consequently, we chose to pursue PACT-sRIMS for whole-mount immunostaining and 3D imaging in this study. However, we note that this is purely a hardware issue, and PACT-RC may be useful with more specialised imaging objectives [18].

PACT-sRIMS permitted imaging of surface structures (Fig. 2b) at marginally improved depths over uncleared tissue (Additional file 4: Figure S4). Using PACT-sRIMS combined with 3D de-noising algorithms (Additional file 5: Figure S5) [25], we were able to visualise surface epithelial structures in 3D at high cellular resolution (Fig. 2b, c and Additional file 6: Figure S6). Figure legends provide a link to higher resolution files. Lactating tissue was particularly amenable to 3D imaging, due to the lower content of adipocytes and the increased surface epithelial mass (Fig. 2c). Using this approach, we were able to observe K5-expressing and SMA-expressing basal cells and K8-expressing and E-cadherin-expressing luminal cells, with confidence.

CUBIC tissue clearing and 3D imaging of the unsectioned mouse mammary gland

Like PACT-sRIMS, CUBIC clearing allowed visualisation of the virgin and lactating mammary epithelia at high cellular resolution (Fig. 3, Additional file 7: Movie 1 and Additional file 8: Figure S7). Additionally, due to the high degree of optical transparency achieved by CUBIC clearing, deeper structures could be readily visualised and imaged using this method (Fig. 3). Whilst we observed strong immunostaining with K5 and SMA antibodies using the CUBIC method, K8 and E-cadherin were not readily and uniformly observed in these conditions. This could be due to sub-optimal fixation or a result of differential protein loss caused by exposure to high levels of detergent without prior sub-cellular stabilisation using a hydrogel monomer. However, using a recent modification to the CUBIC Reagent 1 formulation (Reagent 1A, see "Methods"), we observed improved immunostaining of K8 and E-cadherin (Additional file 9: Figure S8), suggesting that epitope availability is better preserved with this new reagent. The stability of various genetically encoded fluorescent proteins (FPs), including GFP, YFP and RFP, was also assessed and found to be adequately preserved through CUBIC processing (Additional file 10: Figure S9a, b).

We determined that CUBIC-cleared samples were amenable to rehydration, paraffin embedding and immunohistochemical staining following whole-mount imaging (Additional file 10: Figure S9c), facilitating the subsequent 2D cellular analysis of whole-mount-imaged mammary structures. On account of the high level of optical clarity achieved using CUBIC-based tissue clearing and its compatibility with subsequent 2D imaging, we also evaluated its utility for whole-mount histochemical analysis (Fig. 4). Currently, carmine is the most prevalent histochemical stain for assessing mammary gland morphogenesis in whole mount. This pigment stains the mammary epithelia an intense pink/red colour and relies on the solvent xylene or methyl salicylate for subsequent optical transparency [29, 30]. Here, we developed and optimised an alternative histochemical stain for the mammary gland, using the cationic dye methyl green and CUBIC tissue clearing. Using this approach we were able to achieve delicate green/blue staining of epithelial structures in the mammary gland of virgin and lactating mice (Fig. 4a). This new staining approach offers improved colour palette flexibility for dual-

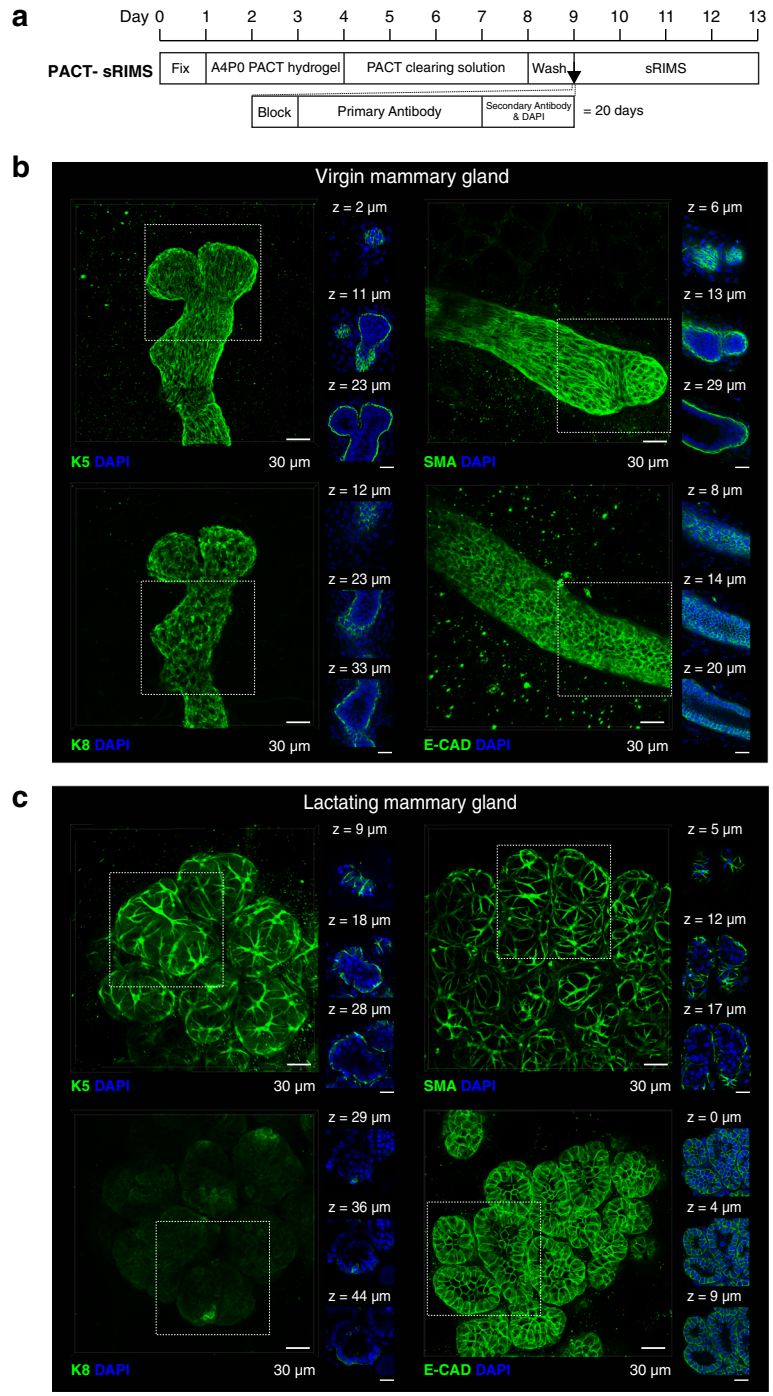


Fig. 2 Passive clarity technique (*PACT*)-sorbitol refractive index matching solution (*sRIMS*) clearing and 3D imaging of virgin and lactating mouse mammary tissue. **a** PACT-sRIMS tissue clearing and immunostaining protocol and timeline. Three-dimensional confocal imaging of PACT-sRIMS-cleared virgin (**b**) and lactating (**c**) mammary glands immunostained with basal cell markers (K5 and smooth muscle actin (*SMA*)) and luminal cell markers (K8 and E-cadherin (*E-CAD*)). *Main image* shows the maximum intensity projection of the entire image sequence, with thin optical slices (1 µm) and their depth (*z* value) relative to the first image in the image sequence. These images are representative of images from at least two mice; further examples of PACT-sRIMS-cleared tissue are shown in Additional file 6: Figure S6). *DAPI* 4′,6-diamidino-2-phenylindole. See additional file 18 for a high resolution version of these PDFs and http://rdcu.be/lT3Z for additional high resolution examples of these imaging techniques [38]

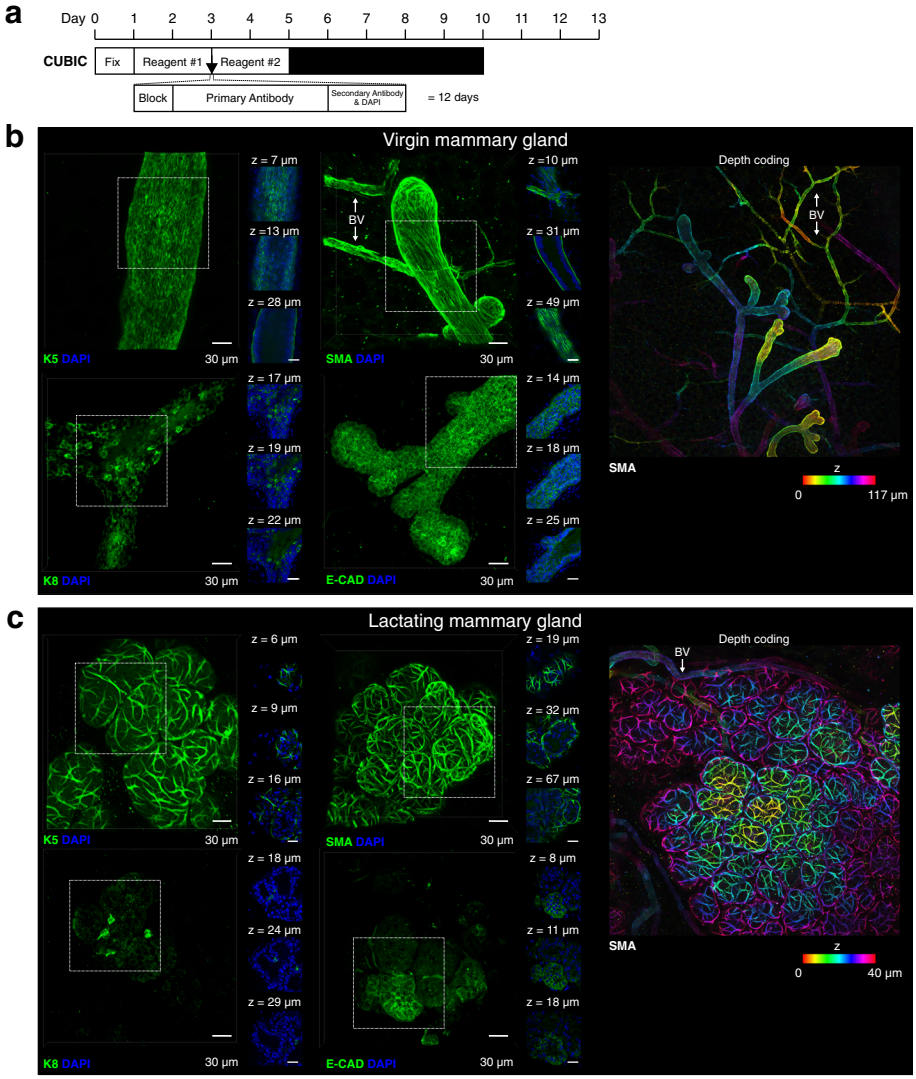


Fig. 3 Clear unobstructed brain imaging cocktails (CUBIC) clearing and 3D imaging of virgin and lactating mouse mammary tissue. **a** CUBIC tissue clearing and immunostaining protocol and timeline. Three-dimensional confocal imaging of CUBIC-cleared virgin (**b**) and lactating (**c**) mammary glands immunostained with basal cell markers (K5 and smooth muscle actin (*SMA*)) and luminal cell markers (K8 and E-cadherin (*E-CAD*)). *Main image* (*green*) shows the maximum intensity projection of the entire image sequence, with thin optical slices (1 µm) and their depth (*z* value) relative to the first image in the image sequence. *Right panel* shows depth-coding of SMA-expressing cells; images in an image stack are assigned a colour based on their relative depth. These images are representative of images from more than three mice; further examples of CUBIC-cleared tissue are shown in Additional file 8: Figure S7 and a modified (Reagent 1A) CUBIC protocol in Additional file 9: Figure S8). *BV* blood vessel (SMA-expressing). *DAPI* 4′,6-diamidino-2-phenylindole. See additional file 18 for a high resolution version of these PDFs and http://rdcu.be/lT3Z for additional high resolution examples of these imaging techniques [38]

colour staining, for example with the novel magenta β-glucosidase (SYNbglA) reporter [19] (Fig. 4b). CUBIC clearing was also semi-compatible with the more traditional whole-mount histochemical stains carmine and haematoxylin (Additional file 11: Figure S10), although these were not the focus of this study and staining may be improved with further optimisation. Importantly, CUBIC clearing was compatible with whole-mount DAB immunohistochemical analysis (Fig. 4c), which has not previously been achieved in the mammary gland.

SeeDB tissue clearing for 3D imaging of the unsectioned mouse mammary gland

Similar to CUBIC clearing, mammary ducts and alveoli could be imaged at high cellular resolution and at considerable depths using SeeDB clearing (Fig. 5, Additional file 12: Movie 2 and Additional file 13: Figure S11). SeeDB-cleared mammary tissue was highly compatible with all antibodies (Fig. 5) and endogenous fluorochromes tested (Additional file 14: Figure S12a, b), with structural morphology well preserved by this simple

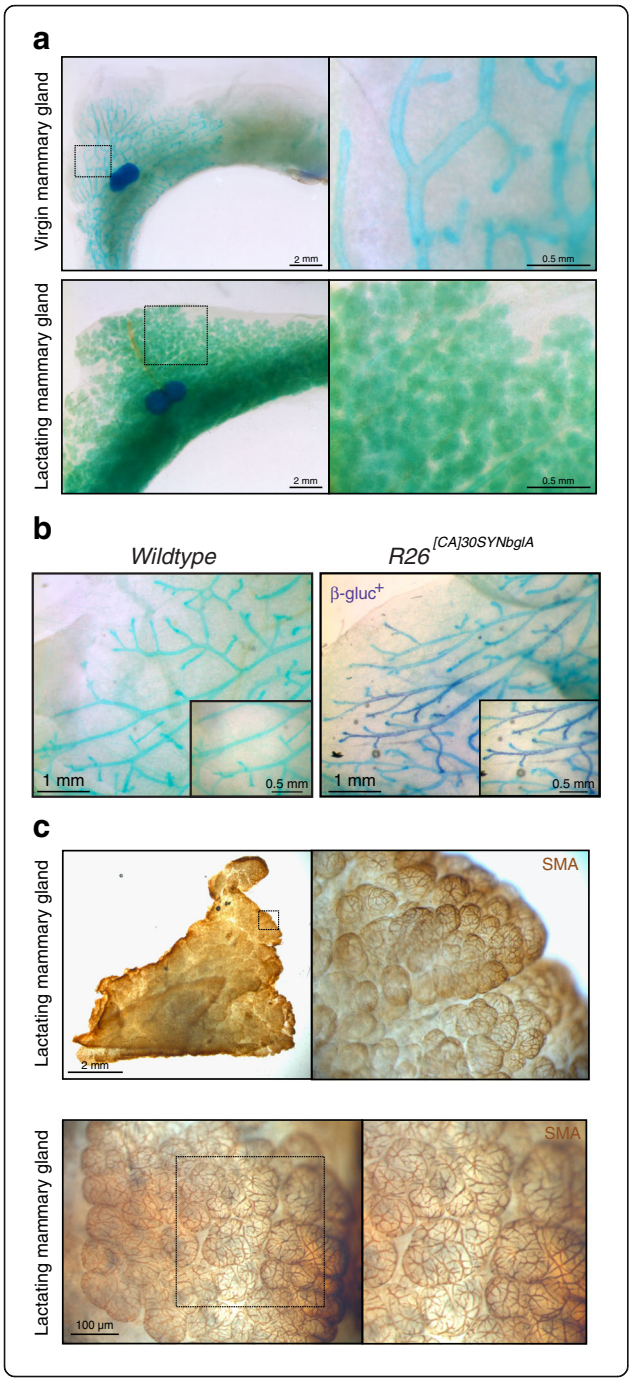


Fig. 4 Clear unobstructed brain imaging cocktails (*CUBIC*) clearing for whole-mount transmission imaging of the mouse mammary gland. **a** Virgin and lactating mammary glands stained with methyl green and cleared with CUBIC for whole-mount morphometric analysis. These images are representative of images from more than three mice. **b** Methyl green counterstaining, showing the compatibility of this light green counterstain with magenta-glu detection of β-glucosidase^{+} cells; β-glucosidase^{+} cells are interspersed with unlabelled cells in this *R26*$^{[CA]30SYNbglA}$ mouse model. **c** Compatibility of CUBIC clearing with smooth muscle actin (*SMA*)-immunostaining and horseradish peroxidase-3,3-diaminobenzidine detection. Immunostaining steps can be performed before CUBIC clearing (*top panel*) or after CUBIC clearing (*bottom panel*). See additional file 18 for a high resolution version of these PDFs and http://rdcu.be/lT3Z for additional high resolution examples of these imaging techniques [38]

fructose-based clearing technique. Additionally, SeeDB-cleared samples were also highly compatible with rehydration and additional whole-mount immunostaining or with paraffin embedding, sectioning and 2D immunostaining (Additional file 14: Figure S12c). Collectively, these properties of optical transparency, morphology preservation, quality of whole-mount immunostaining and suitability for rehydration and 2D analyses, made SeeDB our preferred method for 3D fluorescence imaging of the mammary gland.

Three-dimensional imaging of mammary tumours

We also tested the performance of PACT-sRIMS, CUBIC and SeeDB clearing protocols on mouse mammary tumours derived from the syngeneic TUBO cell line [17]. All clearing protocols permitted high resolution imaging of HER2-expressing cells in mammary tumours at enhanced depths (Fig. 6 and Additional file 15: Movie 3). Similar to virgin and lactating tissue, K8 immunostaining was less intense in CUBIC-cleared mammary tumours (Fig. 6b), but may be improved by the second generation formulation (Reagent 1A). We observed that HER2, K8 and DAPI fluorescence intensity was reduced with increasing imaging depth in all protocols (Fig. 6). Whilst this may in part be attributable to a technical artefact, e.g. sub-optimal fixation or inadequate antibody penetration, it may also be a reflection of the inherent biology and heterogeneity of these tumour samples. Indeed, 2D immunohistochemical analysis revealed that the centre of many tumour lobules contained cleaved caspase-3-positive cells within areas of low E-cadherin staining (Additional file 16: Figure S13), highlighting the value of performing 2D analyses and 3D imaging in parallel when characterising complex and heterogeneous specimens.

Finally, we utilised two-photon excitation microscopy (TPEM) and light sheet fluorescence microscopy (LSFM) for rapid, large-scale imaging of normal and tumourigenic mammary tissue at improved depths and speeds

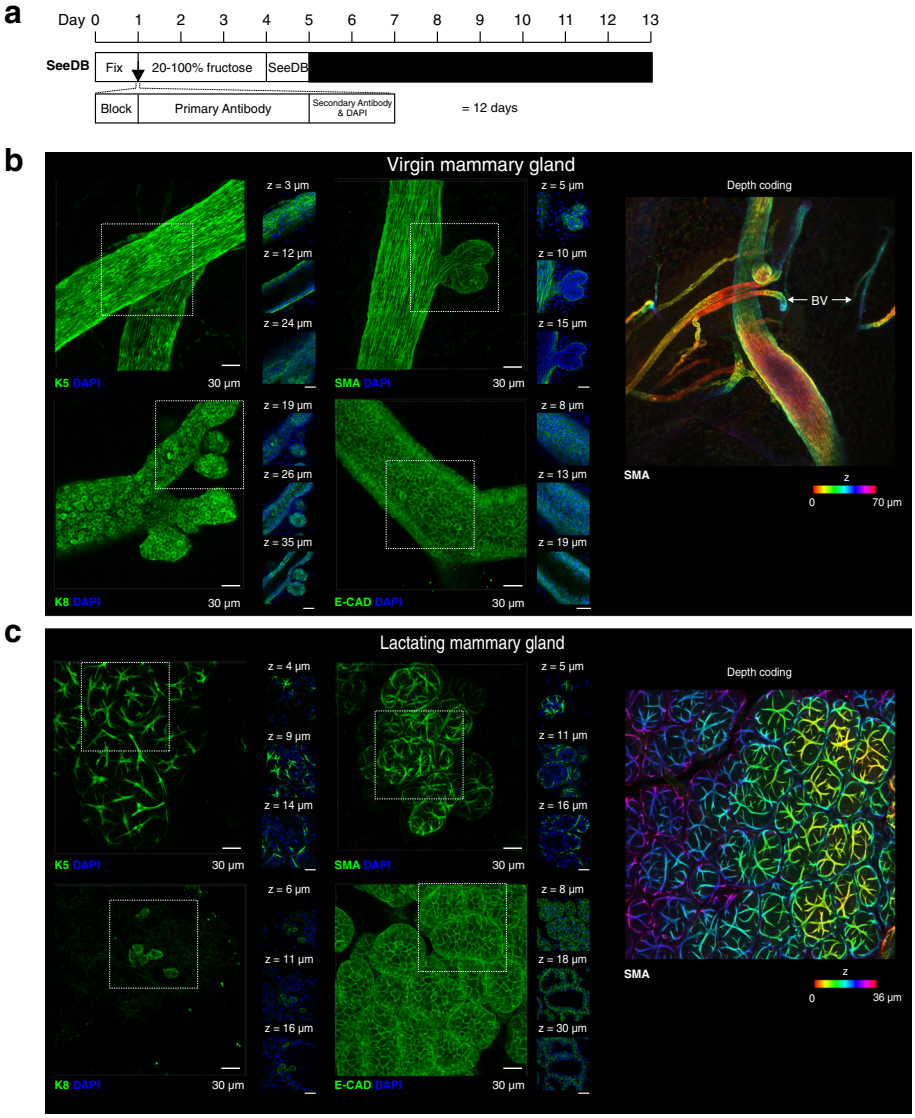


Fig. 5 See deep brain (*SeeDB*)-clearing and 3D imaging of virgin and lactating mouse mammary tissue. **a** SeeDB tissue clearing and immunostaining protocol and timeline. Three-dimensional confocal imaging of SeeDB-cleared virgin (**b**) and lactating (**c**) mammary glands immunostained with basal cell markers (K5 and smooth muscle actin (*SMA*)) and luminal cell markers (K8 and E-cadherin (*E-CAD*)). *Main image* (*green*) shows the maximum intensity projection of the entire image sequence, with thin optical slices (1 µm) and their depth (*z* value) relative to the first image in the image sequence. *Right panel* shows depth-coding of SMA-expressing cells; images in an image stack are assigned a colour based on their relative depth. These images are representative of images from more than three mice; further examples of SeeDB-cleared tissue are shown in Additional file 13: Figure S11). *BV* blood vessel (SMA-expressing). *DAPI* 4′,6-diamidino-2-phenylindole. See additional file 18 for a high resolution version of these PDFs and http://rdcu.be/lT3Z for additional high resolution examples of these imaging techniques [38]

(Fig. 7 and Additional file 17: Figure S14). TPEM uses longer wavelengths, thus allowing deeper penetration, less scattering of light and reduced out-of-focus photobleaching than laser scanning confocal microscopy [31]. Using this technique, we were able to image approx. $1.2 \times 1.2 \times 0.1$ mm of SeeDB-cleared tumour tissue in less than 15 minutes per individual channel at high cellular resolution (Fig. 7). Deeper imaging was achievable (Fig. 7, orthogonal projections); however, as seen with confocal microscopy, HER2 immunostaining declined with increasing depth (Fig. 7, optical slices). Furthermore, using a home-built LSFM [27], with dual side illumination, we were able to image a volume of $0.8 \times 0.8 \times 1.5$ mm in normal mammary tissue with CUBIC clearing in less than 5 min per individual channel (Additional file 17: Figure S14). These data demonstrate that SeeDB-cleared and CUBIC-cleared mammary tissue are also compatible with advanced imaging techniques.

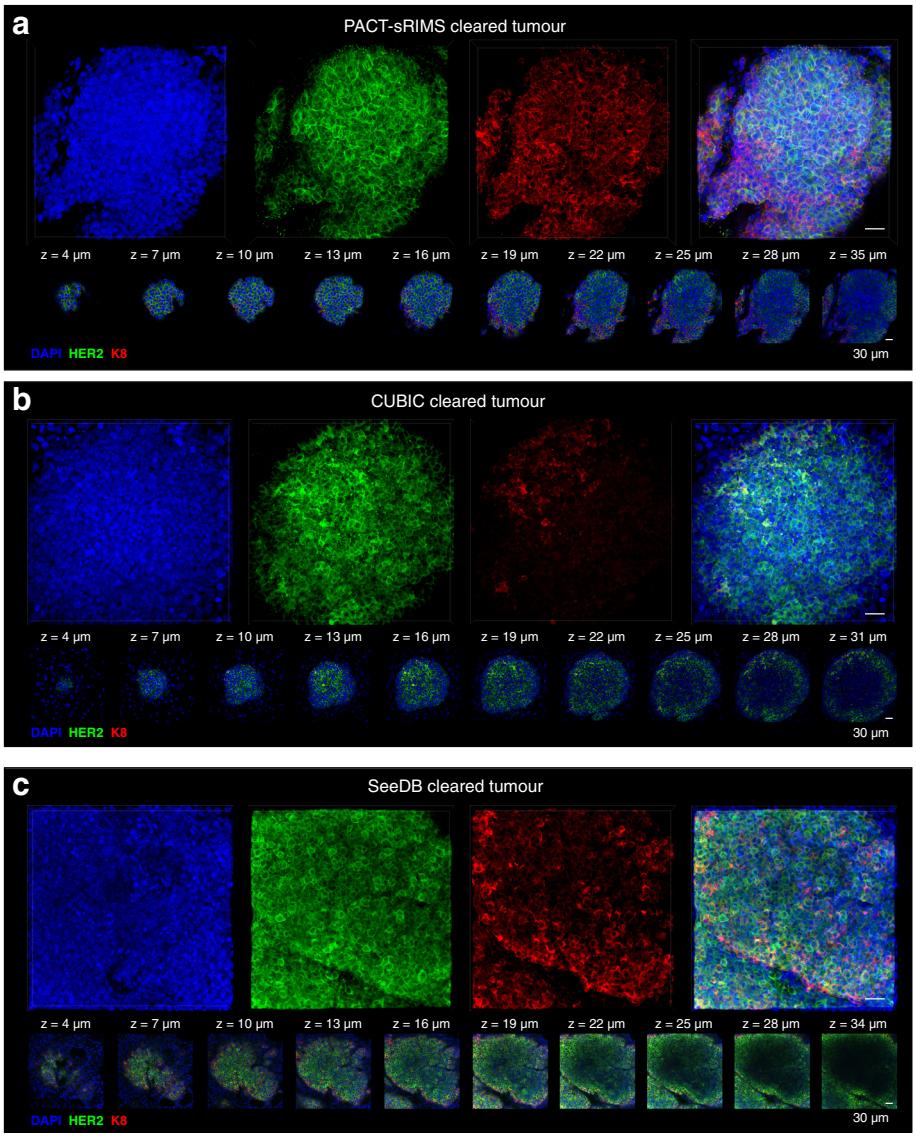


Fig. 6 Three-dimensional confocal imaging of mouse mammary tumours cleared with the passive clarity technique (*PACT*)-sorbitol refractive index matching solution (*sRIMS*), clear unobstructed brain imaging cocktails (*CUBIC*) and see deep brain (*SeeDB*) methods. PACT-sRIMS cleared tumour tissue (**a**), CUBIC-cleared tumour tissue (**b**) and SeeDB-cleared tumour tissue (**c**). Images show maximum intensity projections of 4′,6-diamidino-2-phenylindole (*DAPI*) nuclear staining and human epidermal growth factor receptor 2 (*HER2*) and K8 immunostaining, with thin optical slices (1 μm) and their depth (*z* value) relative to the first image in the image sequence. These images are representative of at least two regions acquired. See additional file 18 for a high resolution version of these PDFs and http://rdcu.be/IT3Z for additional high resolution examples of these imaging techniques [38]

Discussion

Recent developments and refinements in whole-organ tissue clearing and 3D imaging techniques, such as LSFM, optical projection tomography (OPT) and X-ray tomography, have provided unprecedented optical access to intact mammalian tissues [32]. These protocols have been utilised for a range of biological applications, including neuronal circuit reconstruction [8–11, 15, 28, 33], characterisation of arterial wall structure [34] and single-cell lineage tracing in the embryonic heart [35]. In this technical report, we compared leading tissue clearing protocols for optical transparency, structural preservation and 3D fluorescent imaging in the intact mammary gland. This is the first study to employ tissue clearing protocols to examine mammary epithelial cells at single-cell resolution within their native stroma.

We first assessed these protocols for their ability to render opaque mammary tissue transparent. SeeDB and CUBIC clearing techniques achieved the highest degree of optical clarity in mammary tissue (Table 1). Due to

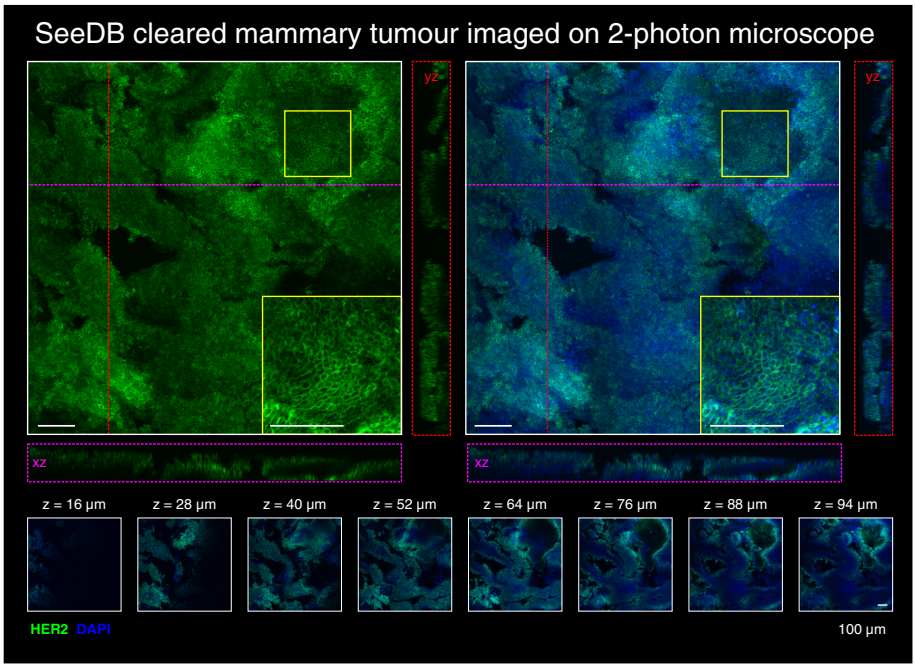


Fig. 7 Imaging of see deep brain (*SeeDB*)-cleared mammary tumours using 2-photon excitation microscopy. Two-photon imaging of SeeDB-cleared human epidermal growth factor receptor 2 (*HER2*)-positive mammary tumours immunostained with HER2 (*green*) and 4′,6-diamidino-2-phenylindole (*DAPI*) nuclear staining (*blue*) and magnified view (*inset*). Orthagonal views show *XZ* (*purple line and box*) and *YZ* (*red line and box*) planes. Thin optical slices (2 µm) and their depth (*z* value) relative to the first image in the image sequence are also shown. See additional file 18 for a high resolution version of these PDFs and http://rdcu.be/IT3Z for additional high resolution examples of these imaging techniques [38]

the exceptional clarity achieved using the CUBIC protocol, we also investigated its suitability as an alternative to xylene delipidation for macroscopic imaging of mouse mammary gland whole mounts. CUBIC clearing combined with methyl green staining resulted in uniform, single-colour staining of mammary epithelial ducts and alveoli, and improved-contrast counterstaining for multi-colour histochemical analysis. Importantly, CUBIC clearing was also compatible with whole-mount chromogenic immunostaining using HRP-DAB detection. Additionally, whole-mount-imaged tissue could be easily recovered and sectioned for subsequent 2D immunohistochemical

Table 1 A comparison of selected clearing methods in the mammary gland

Method	Method overview	RI	Clearing capability	Duration[a]	Preservation		Whole mount IHC	Long-term storage	Rehydration and sectioning	Reference
					structure	FP				
Uncleared	No clearing protocol applied	–	–	1 day	Preserved	Preserved	Compatible	No	Possible[f]	–
3DISCO	Organic solvent-based	1.56	Strong	2 days	Compromised	Rapid loss	Difficult[c]	No	Not possible	[8]
PACT-RC	Aqueous solution-based (hydrogel embedding)	1.45	Weak	10 days	Preserved, mild expansion	Preserved	Compatible	Yes	Not possible	[11]
PACT-sRIMS	Aqueous solution-based (hydrogel embedding)	1.46	Weak	13 days	Preserved	Preserved	Compatible	No	Possible[f]	[11]
CUBIC	Aqueous solution-based (simple immersion)	1.48–9	Strong	5 days	Preserved, mild expansion	Some loss[b]	Semi-compatible[d]	No	Confirmed	[10]
SeeDB	Aqueous solution-based (simple immersion)	1.49	Moderate	5 days	Preserved, mild shrinkage	Preserved	Compatible	No[e]	Confirmed	[9]

[a]Duration from the time of tissue harvest (includes fixation time typically 6–16 h for mammary tissue). [b]As previously reported [10, 39] and the observed requirement for slightly higher laser power for confocal imaging. This may be improved by using the second generation protocol (using Reagent 1A). [c]Using the three-dimensional imaging of solvent-cleared organs combined with optimised whole-mount immunolabelling procedures iDISCO/three-dimensional imaging of solvent-cleared organs (3DISCO) method. The fluorescence signal is rapidly quenched using benzyl alcohol benzyl benzoate (BABB) and a specialised imaging chamber is required for dibenzyl ether (DBE). [d]Some antibodies (i.e. E-cadherin, K8 and human epidermal growth factor receptor 2) do not perform as well as in other clearing protocols. [e]Optimally imaged within 2 weeks but may be stored for several months. [f]Likely to be compatible, but not tested in this study. *RI* refractive index, *FP* fluorescent protein, *IHC* immunohistochemical analysis, *PACT-RC* passive clarity technique-Rapiclear, *sRIMS* sorbitol refractive index matching solution, *CUBIC* clear unobstructed brain imaging cocktails, *SeeDB* see deep brain

analysis. For these reasons, we propose CUBIC clearing as a novel, superior clearing agent for mammary gland whole mounts.

A major limitation of early tissue-clearing techniques is pronounced sample volume changes, leading to cellular deformations [12, 28]. Here, we found that mammary tissue cleared with PACT-RC and CUBIC was associated with a moderate degree of sample expansion, as previously observed in brain tissue cleared using these techniques [10, 11]. Conversely, SeeDB was associated with minor shrinkage of mammary tissue (Table 1), while PACT-sRIMS had no effect on sample volume. Importantly, morphological deformations were not observed by confocal microscopy for any of the four main tissue clearing protocols assessed in this study. Thus, these small, predictable changes in linear sample volume are not likely to pose a major problem for the majority of studies in the mammary gland, provided the control/comparator mammary gland is also subjected to the same clearing protocol. Additionally, in contrast to mechanical dissection or enzymatic digestion, most optical tissue clearing protocols preserve both tissue and matrix architecture, and thus facilitate 3D imaging and analysis of epithelial-stromal interactions.

Whilst important biological information can be garnered from 3D imaging of near-surface structures in uncleared tissue [36, 37], the advantages presented by tissue clearing for the visualisation of expansive areas of mammary tissue are considerable, e.g. for comprehensive clonal analysis in lineage tracing studies [38] or mapping the cellular circuitry driving mammary gland development. Using tissue clearing protocols we were able to image genetically encoded fluorescent proteins and a range of immunolabelled lineage markers at vastly improved depths and at high cellular resolution in virgin, lactating and tumour tissue. Adoption of these protocols combined with volumetric imaging and published denoising algorithms will greatly enhance our understanding of the structural organisation and development of the normal mammary gland, and how these processes are subverted in cancer. An evaluation of the performance of these protocols for diagnostic and experimental studies using clinical tumour biopsies is an aim for the future.

Although not evaluated in this study, it is expected that other antibodies, with appropriate optimisation, would be compatible with whole-mount immunostaining, tissue clearing and 3D imaging in the mammary gland. Thus, this technique has widespread applications in the broader fields of mammary gland biology and pathology. In particular, appropriate optimisation of the tissue fixation time is paramount for whole-mount immunostaining, which is not compatible with heat-induced epitope retrieval. Reduced performance of some antibodies was observed with CUBIC clearing in the mammary gland. This method relies on high concentrations of detergent (15% triton-X100) for clearing and, unlike the PACT protocol, does not entail prior hydrogel embedding to stabilise cellular structures, raising concerns that this protocol may be associated with some protein loss (Table 1) [10]. Indeed, this may explain the compromised fluorescence immunostaining of K8, E-cadherin and HER2 observed with the CUBIC clearing protocol in our study. To overcome this issue an updated CUBIC protocol has recently been developed (unpublished, see "Methods") aimed at improving issues related to protein loss [39]. Alternatively, samples could be gel-embedded according to the CLARITY protocol [15] prior to CUBIC clearing; however, this has not been rigorously assessed in this or other [10] studies, and would require further optimisation for mammary tissue.

An assessment of the qualities of optical transparency, structural preservation, imaging depth, immunostaining, compatibility with downstream analyses, cost and safety (Table 1), placed SeeDB as our method of choice for tissue clearing and volumetric imaging of mammary tissue using standard confocal and advanced fluorescence imaging techniques. The deep tissue imaging of normal and pathological mammary tissue will greatly improve our understanding of this architecturally complex and heterogeneous organ.

Conclusions

This technical report compared the strengths and limitations of a range of whole-organ tissue clearing protocols for optical transparency and 3D imaging in the mouse mammary gland. Notably, the methods examined here are open source protocols, which utilise reagents that are both widely available and affordable to most laboratories. Additionally, whilst these protocols are compatible with advanced imaging techniques, they can also be paired with standard confocal microscopy and open source analysis platforms for universal use. We hope that this publication sheds light on the methods available for optical tissue clearing of the mouse mammary gland, and encourages other researchers to perform their mammary tissue imaging in 3D, with all architectural information preserved.

Additional files

Additional file 1: Figure S1. No primary antibody controls of CUBIC cleared tissue (*top panel*) and SeeDB cleared tissue (*bottom panel*). See Additional file 18 for a high resolution version of these PDFs. (PDF 243 mb)

Additional file 2: Figure S2. 3DISCO clearing and 3D imaging of virgin and lactating mouse mammary tissue. **a** 3DISCO tissue clearing and immunostaining protocol and timeline. **b** Transmission images of 3DISCO cleared tissue. **c** Volume changes caused by 3DISCO-based clearing of virgin and lactating mammary tissue. Values are representative of measurements from three tissue pieces at each developmental timepoint.

d 3D confocal imaging of 3DISCO-cleared virgin and lactating mammary glands immunostained with the basal cellb marker SMA. These images are representative of images from more than two mice. See Additional file 18 for a high resolution version of these PDFs. (PDF 8 mb)

Additional file 3: Figure S3. PACT-RC clearing and 3D imaging of virgin and lactating mouse mammary tissue. **a** PACT-RC tissue clearing and immunostaining protocol and timeline. **b** The optimal (*i*) and PACT-RC (*ii*) imaging configuration for confocal microscopy. For optimal illumination, samples are squeezed between two glass coverslips and are easily flipped or re-positioned for imaging. In contrast, for PACT-RC-cleared samples, samples are mounted using an iSpacer chamber and are difficult to adjust or reposition for optimal illumination, thus the working distance becomes the limiting factor in image acquisition. To overcome this, sample thickness must be closely matched to the thickness of the iSpacer chamber, specialised imaging objectives need to be used or different RI-matching solutions are needed. **c** 3D confocal imaging of PACT-RC-cleared virgin and lactating mammary glands immunostained with the basal cell marker SMA and stained with the nuclear stain DAPI. These images are representative of images from at least two mice. See Additional file 18 for a high resolution version of these PDFs. (PDF 124 mb)

Additional file 4: Figure S4. 3D imaging of uncleared virgin and lactating mouse mammary tissue. **a** Immunostaining protocol and timeline, without a tissue clearing step. 3D confocal imaging of uncleared virgin (**b**) and lactating (**c**) mammary glands immunostained with basal cell marker SMA and the luminal cell marker E-cadherin. Main image shows the maximum intensity projection of the entire image sequence, with thin optical slices (1 μm) and their depth (*z* value) relative to the first image in the image sequence. Visible structures were located very close to the surface of the tissue. See Additional file 18 for a high resolution version of these PDFs. (PDF 176 mb)

Additional file 5: Figure S5. De-noising of 3D image sequences. DAPI staining (*top panel*) and E-cadherin immunostaining (*bottom panel*) in lactating mammary tissue before and after a de-noising algorithm was applied to minimise Poisson-Gaussian noise in the 3D image stacks. De-noising does not greatly alter the outward appearance of the image sequence, but rather aids downstream computer-assisted analyses. See Additional file 18 for a high resolution version of these PDFs. (PDF 4 mb)

Additional file 6: Figure S6. Additional 3D confocal images of PACT-sRIMS-cleared mammary glands, related to Fig. 2. *BV* blood vessel (SMA-expressing). See Additional file 18 for a high resolution version of these PDFs. (PDF 13 mb)

Additional file 7: Movie 1. Three-dimensional rendering of SMA immunostaining (*red*) and DAPI nuclear staining (*blue*) in CUBIC-cleared lactating mammary tissue. Movie shows a layer of basket-like basal cells surrounding each alveolus with single cell resolution. (AVI 20400 kb)

Additional file 8: Figure S7. Additional 3D confocal images of CUBIC-cleared mammary glands, related to Fig. 3. *Arrowhead* shows non-specific intraluminal staining occasionally observed with CUBIC clearing, which may be improved with more rigorous washing following immersion in Reagent 1. See Additional file 18 for a high resolution version of these PDFs. (PDF 23 mb)

Additional file 9: Figure S8. Three-dimensional confocal images of mammary glands cleared using a modified CUBIC clearing protocol (Reagent 1A) to minimise protein loss. Images show the maximum intensity projection of the entire image sequence. *Bottom right panel* shows depth-coding of SMA-expressing myoepithelial cells in lactating mammary tissue; images in an image stack are assigned a colour based on their relative depth. *Arrowhead* shows non-specific intraluminal staining occasionally observed with modified CUBIC clearing, which may be improved with further washing. These images are representative of images from two mice. *BV* blood vessel (SMA-expressing). See Additional file 18 for a high resolution version of these PDFs. (PDF 12 mb)

Additional file 10: Figure S9. Compatibility of CUBIC clearing with genetically encoded FPs and recovery for histological sections. **a** CUBIC clearing of tissue from R26-Confetti mice with reporter expression (nuclear GFP, cytosolic YFP and cytosolic RFP) induced at very low, sporadic levels. Membranous CFP was not observed with any clearing protocol and may be technical; however, this FP is also reportedly underrepresented in mammary tissue from R26-Confetti mice. **b** Tissue from lactating R26-Tdtomato mice induced at very high levels. **c** CUBIC-cleared tissue was rehydrated in PBS prior to standard processing, paraffin embedding and sectioning. Immunostaining for E-cadherin (luminal) and SMA (basal) markers confirm the compatibility of CUBIC clearing with tissue recovery and immunostaining. See Additional file 18 for a high resolution version of these PDFs. (PDF 126 mb)

Additional file 11: Figure S10. Compatibility of CUBIC clearing with carmine and haematoxylin whole-mount staining. Carmine staining (*top*) shows light pink, non-uniform staining in virgin and lactating tissue. Haematoxylin staining (*bottom*) was an intense blue colour in both ducts and stroma. See Additional file 18 for a high resolution version of these PDFs. (PDF 30 mb)

Additional file 12: Movie 2. Three-dimensional rendering of SMA (*green*) and K8 (*red*) immunostaining in SeeDB-cleared virgin mammary tissue. Movie shows luminal and basal cell layers with single cell resolution. (AVI 32300 kb)

Additional file 13: Figure S11. Additional 3D confocal images of SeeDB-cleared mammary glands, related to Fig. 5. See Additional file 18 for a high resolution version of these PDFs. (PDF 10 mb)

Additional file 14: Figure S12. Compatibility of SeeDB clearing with genetically encoded FPs and recovery for histological sections. **a** SeeDB clearing of tissue from virgin R26-Confetti mice with reporter expression (nuclear GFP, cytosolic YFP and cytosolic RFP) induced at very low, sporadic levels. Membranous CFP was not observed with any clearing protocol and may be technical; however, this FP is also reportedly underrepresented in mammary tissue from R26-Confetti mice. **b** Tissue from lactating R26-Tdtomato mice induced at very high levels. **c** SeeDB-cleared tissue was rehydrated in PBS prior to standard processing, paraffin embedding and sectioning. Immunostaining for K8 (luminal) and SMA (basal) markers confirm the compatibility of SeeDB clearing with tissue recovery and immunostaining. See Additional file 18 for a high resolution version of these PDFs. (PDF 154 mb)

Additional file 15: Movie 3. Three-dimensional rendering of HER2 (*red*) and K8 (*green*) immunostaining with DAPI nuclear staining (*blue*) in SeeDB-cleared mammary tumours. Movie shows optical slices and the single cell resolution, allowing visualisation of the 3D structure of the tumour tissue. (AVI 45300 kb)

Additional file 16: Figure S13. Two dimensional immunohistochemical analysis of tumour fragments. K8 and HER2 (**a**) and E-cadherin and cleaved caspase-3 (CC3) (**b**) immunostaining on formalin-fixed paraffin embedded tissue sections from mouse mammary tumours. CC3-positive cells can be observed close to the tumour boundary and in the centre of the tumour fragment, and thus, are not simply an artefact of sub-optimal tissue fixation. *Dotted line* in *left panel* shows the tissue boundary. DAPI nuclear staining (*blue*). See Additional file 18 for a high resolution version of these PDFs. (PDF 26 mb)

Additional file 17: Figure S14. LSFM imaging in the mammary gland. SMA immunostaining and CUBIC clearing of mammary tissue from (**a**) early gestation (total depth 800 μm) and (**b**) virgin (total depth 500 μm) mammary tissue. *BV* blood vessel (SMA-expressing). See Additional file 18 for a high resolution version of these PDFs. (PDF 9 mb)

Additional file 18: All figures in high resolution. (ZIP 127 MB)

Acknowledgements
We thank the Department of Pathology Biological Services Unit for help with animal work, H. Skelton for help with histology, and Peter Humphreys, Drs Leila Muresan, Kevin O'Holleran and Martin Lenz for imaging support. All imaging was performed at the Cambridge Advanced Imaging Centre, Cambridge Stem Cell Institute and the Medical Research Council Laboratory for Molecular Biology (MRC LMB). We thank Adam Fowle and Steve Scotcher from the MRC LMB mechanical workshop for excellent support, and Dr Nick Barry from the light microscopy facility of the MRC LMB for advice on modifications of the light sheet system. We also thank Drs Ueda and Susaki (RIKEN Quantitative Biology Center) for their assistance with CUBIC-based tissue clearing.

Funding
This work was supported by a grant from the Medical Research Council (MRC) programme grant number MR/J001023/1 (BL-L and CJW). FMD was funded by a National Health and Medical Research Council CJ Martin Biomedical Fellowship (GNT1071074). OBH was funded by a Wellcome Trust PhD Studentship (105377/Z/14/Z). JRH was funded by an MRC research grant number MR/K011014/1. FCL was funded by Cancer Research UK and MP was funded by the MRC-LMB (MC_U105178788).

Availability of data and materials
The authors declare that all data supporting the findings of this study are available in the manuscript and its supplementary information files, or are available from the corresponding authors upon request.

Authors' contributions
FMD and BL-L conceived and designed all experiments, performed the optical clearing assays, confocal imaging and whole-mount histochemical analysis and wrote the manuscript. OBH performed confocal imaging and DAB-HRP whole-mount histochemical analysis and helped to revise the manuscript. JRH generated all tumour tissue for confocal and TPEM studies and helped to revise the manuscript. FCL helped design PACT and CUBIC studies, assisted with PACT tissue clearing and helped to revise the manuscript. MP developed and performed all LSFM studies and helped to revise the manuscript. CJW conceived the experiments and helped to revise the manuscript. All authors read and approved the final manuscript.

Authors' information
Drs Felicity Davis, Bethan Lloyd-Lewis, Jessica Hitchcock and Filipe Lourenco are Postdoctoral Fellows and Olivia Harris is a PhD Student at the University of Cambridge. Dr Mathias Pasche is a Technical Support Officer at the Light Microscopy Facility of the MRC Laboratory of Molecular Biology. Prof Christine Watson is a Group Leader at the University of Cambridge and Vice Principal of Newnham College.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
All animal experimentation was carried out in accordance with the *Animal (Scientific Procedures) Act 1986*, the *European Union Directive 86/609* and with local ethics committee approval (University of Cambridge Animal Welfare and Ethical Review Body (AWERB), UK Home Office and The University of Queensland Animal Ethics Committee).

Author details
[1]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK. [2]School of Pharmacy, The University of Queensland, Brisbane 4072, Australia. [3]Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 1QR, UK. [4]Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, UK. [5]Medical Research Council Laboratory for Molecular Biology, Cambridge CB2 0QH, UK.

Received: 24 June 2016 Accepted: 18 August 2016
Published online: 13 December 2016

References

1. Macias H, Hinck L. Mammary gland development. Wiley Interdiscip Rev Dev Biol. 2012;1:533–57.
2. Rios AC, Fu NY, Lindeman GJ, Visvader JE. In situ identification of bipotent stem cells in the mammary gland. Nature. 2014;506:322–7.
3. Wuidart A, Ousset M, Rulands S, Simons B, Van Keymeulen A, Blanpain C. Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. Genes Dev. 2016;30:1261–77.
4. Wang D, Cai C, Dong X, Yu QC, Zhang X-O, Yang L, et al. Identification of multipotent mammary stem cells by protein C receptor expression. Nature. 2015;517:81–4.
5. Rios AC, Fu NY, Jamieson PR, Pal B, Whitehead L, Nicholas KR, et al. Essential role for a novel population of binucleated mammary epithelial cells in lactation. Nat Commun. 2016;7:11400.
6. Spalteholz W. Uber das Durchsichtigmachen von menschlichen und tierischen Praparaten. Leipzig; 1914.
7. Tainaka K, Kuno A, Kubota S, Murakami T, Ueda H. Chemical principles in tissue clearing and staining protocols for whole-body cell profiling. Annu Rev Cell Dev Biol. 2016;32:1–29.
8. Erturk A, Becker K, Jahrling N, Mauch CP, Hojer CD, Egen JG, et al. Three-dimensional imaging of solvent-cleared organs using 3DISCO. Nat Protoc. 2012;7:1983–95.
9. Ke M-T, Fujimoto S, Imai T. SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. Nat Neurosci. 2013;16:1154–61.
10. Susaki EA, Tainaka K, Perrin D, Kishino F, Tawara T, Watanabe TM, et al. Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. Cell. 2014;157:726–39.
11. Yang B, Treweek JB, Kulkarni RP, Deverman BE, Chen CK, Lubeck E, et al. Single-cell phenotyping within transparent intact tissue through whole-body clearing. Cell. 2014;158:945–58.
12. Becker K, Jährling N, Saghafi S, Weiler R, Dodt HU. Chemical clearing and dehydration of GFP expressing mouse brains. PLoS One. 2012;7:e33916.
13. Dodt H-U, Leischner U, Schierloh A, Jährling N, Mauch CP, Deininger K, et al. Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. Nat Methods. 2007;4:331–6.
14. Renier N, Wu Z, Simon DJ, Yang J, Ariel P, Tessier-Lavigne M. iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. Cell. 2014;159:896–910.
15. Chung K, Deisseroth K. CLARITY for mapping the nervous system. Nat Methods. 2013;10:508–13.
16. Tainaka K, Kubota SI, Suyama TQ, Susaki EA, Perrin D, Ukai-Tadenuma M, et al. Whole-body imaging with single-cell resolution by tissue decolorization. Cell. 2014;159:911–24.
17. Rovero S, Amici A, Di Carlo E, Bei R, Nanni P, Quaglino E, et al. DNA vaccination against rat her-2/Neu p185 more effectively inhibits carcinogenesis than transplantable carcinomas in transgenic BALB/c mice. J Immunol. 2000;165:5133–42.
18. Marx V. Microscopy: seeing through tissue. Nat Methods. 2014;11:1209–14.
19. McCutcheon SC, Jones K, Cumming SA, Kemp R, Ireland-Zecchini H, Saunders JC, et al. Characterization of a heat resistant beta-glucosidase as a new reporter in cells and mice. BMC Biol. 2010;8:89.
20. Manousiouthakis E, Mendez M, Garner MC, Exertier P, Makita T. Venous endothelin guides sympathetic innervation of the developing mouse heart. Nat Commun. 2014;5:3918.
21. Dills WL. Protein fructosylation: fructose and the Maillard reaction. Am J Clin Nutr. 1993;58:779S-87S.
22. Sargeant TJ, Lloyd-Lewis B, Resemann HK, Ramos-Montoya A, Skepper J, Watson CJ. Stat3 controls cell death during mammary gland involution by regulating uptake of milk fat globules and lysosomal membrane permeabilization. Nat Cell Biol. 2014;16:1057–68.
23. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open source platform for biological image analysis. Nat Methods. 2012;9:676–82.
24. Linkert M, Rueden CT, Allan C, Burel JM, Moore W, Patterson A, et al. Metadata matters: access to image data in the real world. J Cell Biol. 2010;189:777–82.
25. Boulanger J, Kervrann C, Bouthemy P, Elbau P, Sibarita J-B, Salamero J. Patch-based nonlocal functional for denoising fluorescence microscopy image sequences. IEEE Trans Med Imaging. 2010;29:442–54.
26. Preibisch S, Saalfeld S, Tomancak P. Globally optimal stitching of tiled 3D microscopic image acquisitions. Bioinformatics. 2009;25:1463–5.
27. Pitrone P, Schindelin J, Stuyvenberg L, Preibisch S, Weber M, Eliceiri K, et al. OpenSPIM: an open-access light-sheet microscopy platform. Nat Methods. 2013;10:598–9.
28. Hama H, Kurokawa H, Kawano H, Ando R, Shimogori T, Noda H, et al. Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. Nat Neurosci. 2011;14:1481–8.
29. Plante I, Stewart MKG, Laird DW. Evaluation of mammary gland development and function in mouse models. J Vis Exp. 2011;53:2–6.
30. van Amerongen R. Lineage tracing in the mammary gland using Cre/lox technology and fluorescent reporter alleles. Methods Mol Biol. 2015;1293:187-211.
31. Drobizhev M, Makarov NS, Tillo SE, Hughes TE, Rebane A. Two-photon absorption properties of fluorescent proteins. Nat Methods. 2011;8:393–9.

32. Shearer T, Bradley R, Hidalgo-Bastida A, Sherratt M, Cartmell S. Three-dimensional visualisation of soft biological structures by X-ray computed micro-tomography. J Cell Sci. 2016;129:2483–92.
33. Schmitt O, Modersitzki J, Heldmann S, Wirtz S, Fischer B. Image registration of sectioned brains. Int J Comput Vis. 2007;73:5–39.
34. Walton LA, Bradley RS, Withers PJ, Newton VL, Watson REB, Austin C, et al. Morphological characterisation of unstained and intact tissue micro-architecture by x-ray computed micro- and nano-tomography. Sci Rep. 2015;5:10074.
35. Li J, Miao L, Shieh D, Spiotto E, Li J, Zhou B, et al. Single-cell lineage tracing reveals that oriented cell division contributes to trabecular morphogenesis and regional specification. Cell Rep. 2016;15:158–70.
36. Davis FM, Janoshazi A, Janardhan KS, Steinckwich N, D'Agostin DM, Petranka JG, et al. Essential role of Orai1 store-operated calcium channels in lactation. Proc Natl Acad Sci U S A. 2015;112:5827–32.
37. Raymond K, Cagnet S, Kreft M, Janssen H, Sonnenberg A, Glukhova MA. Control of mammary myoepithelial cell contractile function by α3β1 integrin signalling. EMBO J. 2011;30:1896–906.
38. Davis, Lloyd-Lewis et al. Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. Nature Communications. 2016. doi:10.1038/ncomms13053.
39. Susaki EA, Ueda HR. Whole-body and whole-organ clearing and imaging techniques with single-cell resolution: toward organism-level systems biology in mammals. Cell Chem Biol. 2016;23:137–57.

CellPress

Review

Mammary Stem Cells: Premise, Properties, and Perspectives

Bethan Lloyd-Lewis,[1,5,*] Olivia B. Harris,[1,2] Christine J. Watson,[1,2] and Felicity M. Davis[3,4,*]

Adult mammary stem cells (MaSCs) drive postnatal organogenesis and remodeling in the mammary gland, and their longevity and potential have important implications for breast cancer. However, despite intense investigation the identity, location, and differentiation potential of MaSCs remain subject to deliberation. The application of genetic lineage-tracing models, combined with quantitative 3D imaging and biophysical methods, has provided new insights into the mammary epithelial hierarchy that challenge classical definitions of MaSC potency and behaviors. We review here recent advances – discussing fundamental unresolved properties of MaSC potency, dynamics, and plasticity – and point to evolving technologies that promise to shed new light on this intractable debate. Elucidation of the physiological mammary differentiation hierarchy is paramount to understanding the complex heterogeneous breast cancer landscape.

Trends

The ability of stem cells to generate and repeatedly regenerate the elaborate mammary epithelium at any stage during the female reproductive lifespan reveals the extraordinary capacity and longevity of stem cells in the adult mammary gland (breast).

Some breast cancers may also be hierarchically organized, containing cancer stem cells that propel and sustain tumor growth and therapy-resistant regrowth.

The adoption and refinement of novel mouse models, imaging techniques, and quantitative methodologies have provided insights into the mammary epithelial cell hierarchy that could not have been achieved using conventional approaches.

Evolving techniques to label, trace, and profile individual cells, both in the normal mammary gland and in tumors, will offer further insights into the capacity, multiplicity, and dynamicity of mammary stem cells and their malignant counterparts.

Adult Mammary Stem Cells: Concepts and Challenges

Adult stem cells exist in diverse organs, such as the intestine, skin, and skeletal muscle [1,2]. In these tissues their primary role is homeostatic, that is, to replenish cells lost to attrition or injury. However, unlike many other organs, the mammary gland primarily develops postnatally [3,4] (Figure 1), and thus stem cells in the adult mammary gland serve both developmental and homeostatic functions.

Construction of the branching ductal epithelium during puberty is driven by hormones, growth factors, and local signaling cues, and proceeds via proliferation of **mammary stem cells** (MaSCs, see Glossary) and their progeny within bulbous distal structures known as **terminal end buds** (TEBs) (Figure 1B) [5,6]. By the end of puberty, ductal morphogenesis is complete and the TEBs have fully regressed (Figure 1C) [4]. Although it is generally accepted that stem cells persist in the adult mammary gland following the demise of the TEBs – where they have essential roles in the generation and regeneration of the alveolar (milk-producing) epithelium during pregnancy and lactation (Figure 1C–F) – the location of these cells within the complex ductal epithelium remains elusive [6,7]. In addition, despite intense investigation and debate, the differentiation potential of adult MaSCs (i.e., their ability to generate one or both of the mammary epithelial cell lineages) remains contentious [7–18]. The longevity and extensive self-renewal properties of these cells, however, place them as probable candidates for oncogenic transformation in some breast cancers [19,20]. Moreover, some breast cancers may be hierarchically organized and contain a pool of cancer stem cells that drive their precipitous long-term growth and regrowth [19–21]. Thus, a greater understanding of the identity, plasticity, and differentiation potential of adult MaSCs, and the specific pathways that regulate their

[1]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK
[2]Wellcome Trust–Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 1QR, UK
[3]School of Pharmacy, University of Queensland, Brisbane 4102, Australia
[4]Mater Research Institute, University of Queensland, Brisbane 4102, Australia

 http://dx.doi.org/10.1016/j.tcb.2017.04.001

self-renewal and fate, may also provide important insights into the heterogeneity and treatment-resistance of this intractable disease.

Recent studies, using single-cell **lineage-tracing** approaches, have revealed the immense capacity of a single MaSC to contribute to the formation of the ductal epithelium during puberty [16,18] and of the alveolar epithelium during pregnancy/lactation [16]. These studies also highlight considerable redundancy within this system [16,18], positing that several hundred lineage-restricted MaSCs actively and stochastically contribute to ductal and alveolar morphogenesis under physiological conditions. This is not entirely surprising, given that lactation is an evolutionarily essential aspect of mammalian survival that demands functional stem cells. However, if MaSCs are the cell of origin in some breast cancers, then this superfluity brings with it a heightened opportunity for oncogenic transformation. Regardless, the inextricable connections between MaSCs and breast cancer warrants further investigation to achieve a unified and enduring characterization of their potential, anatomical location, and molecular profile.

We discuss here recent insights into the **mammary epithelial cell hierarchy**, addressing unanswered questions relating to MaSC potency, dynamics, and plasticity. We discuss the unique challenges in elucidating the mammary epithelial cell hierarchy, and highlight evolving technologies that promise to shed new light on these difficult issues.

The Epithelial Cell Hierarchy: An Evolving Paradigm

In 1959, a seminal study published in *Cancer Research* demonstrated that fragments of mammary tissue could be transplanted into the epithelium-divested fat pad of a recipient

[5]Current address: Department of Genetics and Developmental Biology, Institut Curie, Paris, France

*Correspondence: bethan.lloyd-lewis@curie.fr (B. Lloyd-Lewis) and f.davis@uq.edu.au (F.M. Davis).

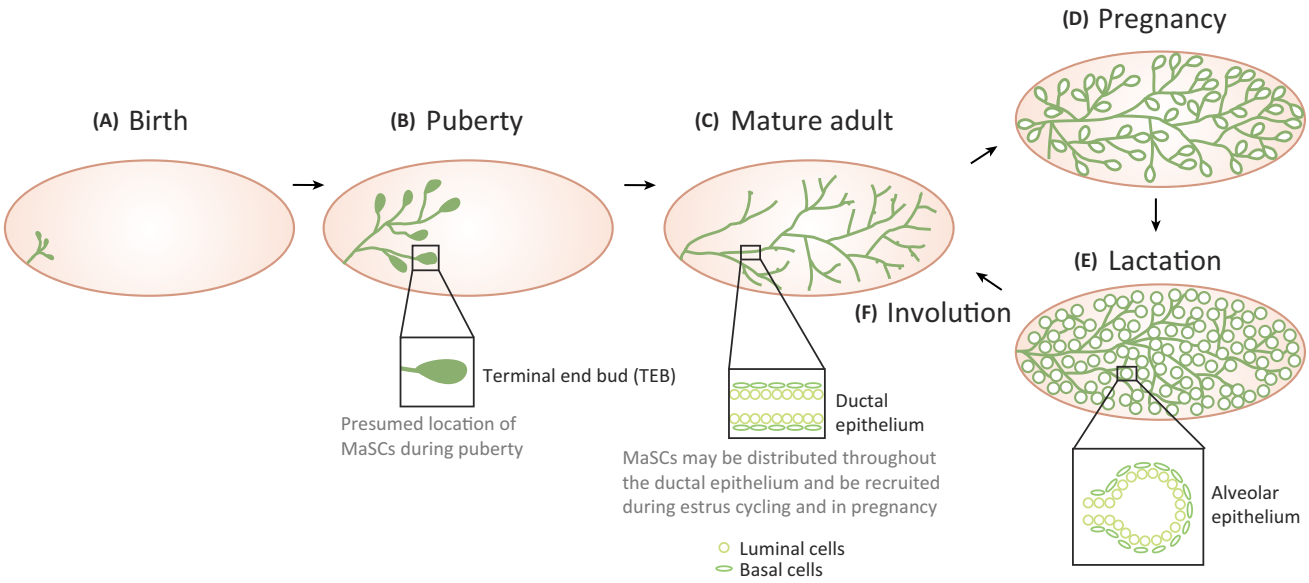


Figure 1. Postnatal Mammary Gland Development in Mice. (A) Mammals are born with only a rudimentary ductal structure (see [3,4] for a description of embryonic mammary gland development) which begins to elongate and invade the empty fat pad at puberty (B). By the end of puberty (C), the ductal structures have reached the boundaries of the mammary fat-pad and the terminal end buds (TEBs) have fully regressed. Mammary ducts are composed of two epithelial cell lineages arranged into distinct cell layers; luminal cells line the lumen of each duct and are surrounded by an outer layer of basal cells (depicted inset). Whether MaSCs in the adult mammary gland are lineage-restricted or can give rise to both luminal and basal cells is an area of contention. (C–E) Resident mammary stem cells (MaSCs) in the mature mammary epithelium are responsible for the generation of milk-producing alveoli during pregnancy and lactation. (F) Stem cells are likely to survive post-lactational regression (involution) to enable successive lactations. The mouse is an excellent model for studying processes regulating human mammary gland development and tumorigenesis, however, key differences exist [83]. Notably, the human mammary gland is arranged in distinct lobes, each with a separate ductal structure and outlet.

mouse, successfully engraft, and generate an entire ductal epithelium anew [22]. What followed was a divisive pursuit to identify and characterize the cells responsible for the development, maintenance, and regeneration of the mammary epithelium (i.e., adult MaSCs) that has lasted for more than 50 years (Figure 2). Transformative advances came on the back of at least three key enabling methodologies: the isolation of cells with enhanced repopulating and self-renewal properties upon transplantation; population-based genetic fate-mapping; and stochastic, single-cell genetic lineage-tracing. The ability to image ducts and alveoli in 3D [14–16,18,23,24] or 4D [18] (Box 1), combined with quantitative image analysis and biostatistical modeling [5,15,18], has also provided important insights into clonal dynamics and dispersion patterns that could not have been attained through the examination of thin tissue sections [25]. We broadly examine these techniques, summarizing key findings with a retrospective wisdom.

Transplantation

The observation that any fragment of mammary tissue has the potential to regenerate the entire bilayered mammary epithelium upon serial transplantation provided strong evidence that **mammary repopulating cells** (believed to be *bona fide* MaSCs) were distributed throughout the length of the adult ductal epithelium [26–28]. Subsequent work using retroviral-tagged mammary tissue fragments [29] and limiting dilutions of heterogeneous cell suspensions confirmed these results [30,31], and these studies were in turn refined and expanded by the identification and purification of a subset of cells with superior repopulating capacity [8–10] (Figure 2). Collectively, these analyses supported the notion that adult MaSCs are

Box 1. Emerging Technologies

Transplantation assays and population-based genetic lineage-tracing studies have largely guided our interrogation of the mammary epithelial cell hierarchy. Now, techniques to label/barcode, visualize, and profile individual cells in real-time in the mammary gland are offering fresh insights into the organization and plasticity of MaSCs and the pathways that dictate their fate. An overview of some of these technologies, many of which have already been utilized by mammary gland biologists and breast cancer researchers, is given below.

Tissue clearing. Tissue clearing techniques [e.g., CUBIC (clear, unobstructed brain imaging cocktail) and SeeDB (see deep brain)] render whole organs transparent and are largely based on minimizing light scattering at oil–water interfaces [73,74]. These techniques facilitate *in situ* 3D imaging of mammary epithelial cells at high cellular resolution [24].

3D and 4D Imaging. Platforms for rapid, multiscale imaging of cells, tissues, and organs include confocal, two-photon, and light sheet microscopy [24]. Tissue clearing can be combined with 3D imaging for the visualization of fixed tissues. Limited 4D imaging can be performed on *ex vivo* tissue [75]; however, these studies are confounded by various artefacts, such as tissue hypoxia, and thus their use is generally restricted to short-term imaging. The surgical insertion of an imaging window for intravital imaging overcomes these issues and allows long-term visualization of cells deep within tissues of living mice [18,76,77].

Single-Cell RNA-Seq. This method allows whole-transcriptome profiling of individual cells [78]. This may be combined with methods to either (i) label single cells and all of their progeny (e.g., single-cell lineage-tracing) [18] or (ii) incrementally mark cells (e.g., CRISPR-generated barcoding), to create a molecular profile of all cells within a single clone.

CRISPR-Generated Barcoding. Genome editing using CRISPR facilitates the interrogation of lineage relationships by creating cumulative genetic mutations (or 'barcodes') in cells without interfering with their normal development or function. Genome editing of synthetic target arrays for lineage tracing (GESTALT) [79] and mammalian synthetic cellular recorders integrating biological events (mSCRIBE) [80] are two such examples. Computational analysis of common mutations allows the reconstruction of lineage relationships.

Ex Vivo Models. 3D culture systems, such as organoid culture [81], facilitate *ex vivo* modeling of epithelial tissue morphogenesis. Using this technology, adult stem cells can be maintained long-term in culture, driving the formation of highly ordered, multicellular structures that recapitulate the organization of their resident tissues. Mammary organoids [82] facilitate the interrogation of epithelial cell–cell interactions, hormone regulation, and paracrine signaling in an accessible, but normal, mammary epithelial cell environment.

Glossary

Basal cell: one of the two main cell lineages in the mammary gland; basal cells surround the luminal cell layer and typically express cytokeratin-5, -14, and smooth muscle actin.
Bi/multipotent: able to give rise to more than one cell lineage, for example a bipotent MaSC may be able to give rise to both basal and luminal progeny.
Clone: all of the progeny of a single parent cell.
Label-retaining cell: a cell that is able to retain a label (be it a lipophilic dye, DNA precursor nucleoside, or regulated expression of a fluorescently tagged histone) over a defined chase period. Cells that remain in cycle dilute the label, whereas slow-cycling or quiescent cells remain labeled at the end of the assay.
Lineage tracing: a technique to identify the progeny of a single cell; the phrase 'population-based lineage-tracing' has been used here to distinguish techniques that trace the progeny of specific populations of cells (e.g., cytokeratin-14-expressing cells) generally at levels higher than clonal density.
Luminal cell: one of the two main cell lineages in the mammary gland; luminal cells line the lumen of ducts and alveoli, they typically express cytokeratin-8, and may be hormone receptor positive or negative.
Mammary epithelial cell hierarchy: the organization of stem, progenitor, and differentiated cells in the mammary gland.
Mammary repopulating cells: cells enriched for the ability to regenerate the mammary epithelium upon serial transplantation at limiting dilution into the cleared fat pad of a recipient mouse.
Mammary stem cells (MaSCs): undifferentiated cells in the mammary gland that are capable of giving rise indefinitely to more stem cells (self-renewal) as well as to more-differentiated daughters through symmetric and asymmetric divisions. Uncertainties surrounding the identity, differentiation potential, and plasticity of these cells have generated semantic debate, and MaSCs are also referred to more conservatively as 'stem/progenitor cells'.

bi/multipotent. The demonstration that lineage-restricted cells could be forced to adopt a multipotent fate under 'regenerative conditions' [11–13] challenged this dogma. It is now widely accepted that mammary repopulating cells, identified by transplantation, are distinct from stem cells that exist under physiological conditions. Nevertheless, this technique has provided some important insights into the qualities of self-renewal and regeneration, with enduring relevance.

Potential stem cell: a more differentiated cell that is able to reacquire stem-like properties under regenerative/wounding conditions; also known as facultative stem cells.
Stem cell niche: the specialized microenvironment in which a stem cell resides that can regulate stem cell self-renewal, differentiation, and longevity.
Terminal end bud (TEB): bulbous proliferative structures at the ends of each main duct during puberty; the presumptive location of pubertal MaSCs.
Unipotent: able to give rise to one main cell lineage; for example, a unipotent luminal stem cell is able to give rise only to luminal progeny, and a unipotent basal stem cell is able to give rise only to basal progeny.

Population-Based Genetic Fate-Mapping

The application of genetic lineage-tracing techniques to mammary tissue has enabled temporal examination of lineage relationships under physiological conditions. These studies have utilized tamoxifen- or doxycycline-responsive transgenic mouse models to induce the expression of reporter genes in predefined cohorts of cells [11–15,17]. The genetic label, typically a fluorescent or histochemical reporter, is permanently expressed by the original cell and is transmitted to all of its progeny. An analysis of reporter expression through time can be used to determine whether the original labeled population contained lineage-restricted stem cells or cells with multi-lineage differentiation potential (Figure 2). In its original application in the mammary gland [11], this approach was used to track the fate of **luminal cells** [e.g., cytokeratin (K)8-expressing] and **basal cells** (e.g., K14-expressing), demonstrating that lineage-restricted MaSCs drive postnatal mammary gland development and maintenance. Subsequent lineage-tracing studies have provided evidence in support of both **unipotent** and bi/multipotent adult MaSCs [12–15,17,32] (Figure 2). Lineage-restricted cell populations have also been shown to convert to multipotency *in vivo* by oncogenic PI3KCA signaling, suggesting that there is scope for plastic transformation, thereby adding further complexity to this system [33,34].

Inconsistencies in recent lineage-tracing studies in the normal mammary gland may be in part attributable to the temporal expression of pathway-specific promoters [12,13] or the fidelity of pan-lineage promoters. Given that a single mammary stem/progenitor cell is capable of producing many hundred progeny [16], the promiscuous labeling of even a small number of cells of the opposing lineage could significantly confound downstream lineage analysis in this model [15,16]. A second limitation relates to the power of population-based labeling approaches to accurately detect the expansion of a single **clone**, which is a function of both the method of detection and the initial labeling density (Figure 3). To overcome this problem, as well as potential tracing artefacts associated with the preferential labeling of specific (and potentially non-representative) cell subpopulations, a recent study has mapped the fate of all basal cells (a technique termed saturation lineage-tracing) [15]. If rare bipotent MaSCs do reside in the basal compartment, and contribute even minimally to mammary gland morphogenesis and homeostasis [1], this could be detected by an increase in the number of fluorescently labeled luminal cells observed using either fluorescence-activated cell sorting or 3D image quantification. No population flux was detected using either method of analysis in these studies, suggesting that basal MaSCs are indeed lineage-restricted [15]. A subsequent report [35], however, demonstrated that enzymatic digestion before 3D imaging [15,18,36,37] can deplete or structurally damage basal cells, postulating that rare bi-lineage clones are not detected under these conditions [35]. Recently described methods for non-proteolytic 3D imaging [16,24], together with quantitative platforms for image analysis which consider tissue architecture, cell morphology, chimerism, and Cre-specificity [15,35], will undoubtedly aid future lineage-tracing studies in the mammary gland.

Stochastic, Single-Cell Genetic Lineage-Tracing

Lineage-tracing has facilitated *in situ* examination of MaSC properties under conditions of minimal interference. However, unlike transplantation assays, these studies have been unable to map the fate of a single labeled cell [9]. Obstacles to single-cell genetic lineage-tracing have,

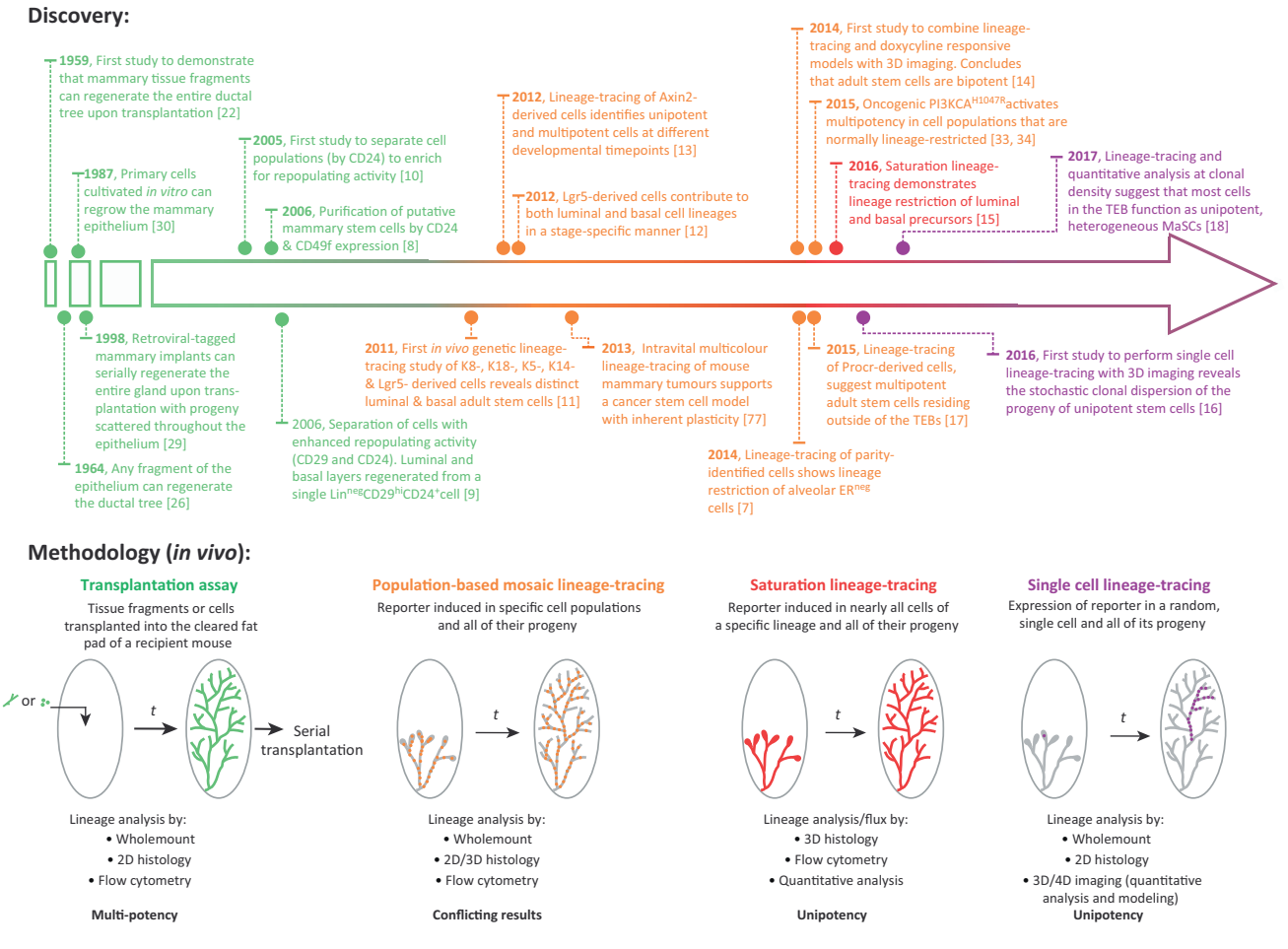


Figure 2. A Summary of the Key Discoveries in the Field and the Methodologies That Enabled These Advances. This timeline focuses on discoveries made within the past decade using transplantation or genetic lineage-tracing assays (see [19,20,84] for more-detailed historical reviews). Schematic diagrams summarizing each *in vivo* methodology are depicted at puberty; however, these techniques have also been utilized to assess cell fate at other developmental stages, and in some cases their use has also been extended to investigate cellular dynamics in mammary tumorigenesis (also see [7–18,22,26,29,30,33,34,77]). Abbreviations: ER, estrogen receptor; MaSC, mammary stem cell; TEB, terminal end bud.

however, been mitigated in part by advances in whole-organ clearing [24] and high-resolution 3D imaging [14] (Box 1).

Recently, R26$^{[CA]30}$ mice [38] have been used to achieve unbiased labeling of single proliferating cells in the mammary gland [16]. Genetic labeling in this model is exceedingly rare, and thus it can be combined with 3D imaging to track the fate of a single labeled cell with confidence (Figure 2). A similar approach to achieve low-density, unbiased labeling involves the use of mice that express inducible Cre recombinase in all cells (R26^{CreERT2}). Neutral, multicolor labeling is achieved by crossing these mice with R26^{Confetti} animals, and sparse reporter induction is attained using low doses of tamoxifen [16,18]. Recent application of these models has provided further evidence that unipotent MaSCs drive ductal morphogenesis during puberty [16,18] and alveolar morphogenesis during gestation [16]. However, although the small number of cells initially labeled in these models permits the indisputable analysis of clonal progeny, it also limits their power to detect and characterize the full spectrum of stem and progenitor cells that are

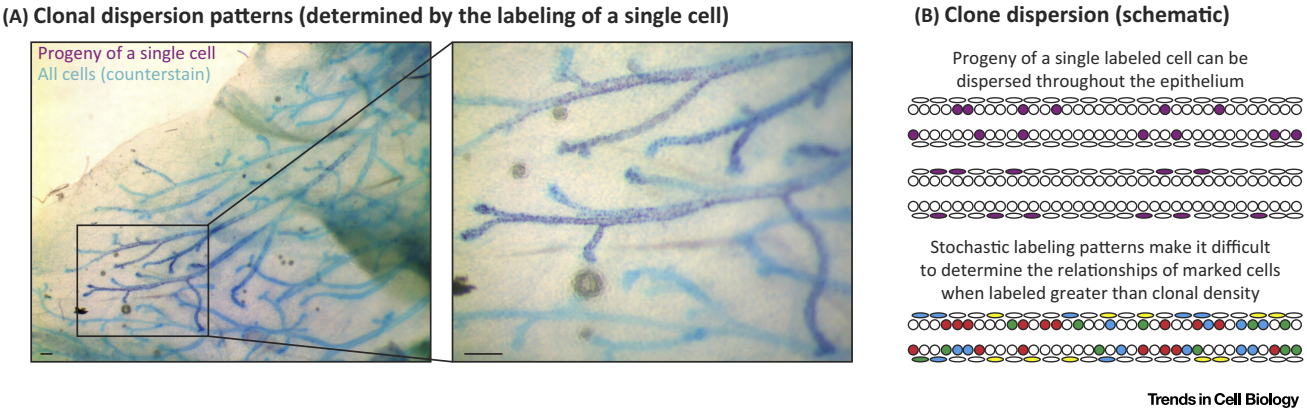


Figure 3. Limitations of Population-Based Lineage-Tracing Studies. (A) Clonal patterns arising from the genetic labeling of a single cell (purple). These studies demonstrate that the progeny of a single marked cell can be distributed throughout the length of the ductal epithelium in a stochastic, interspersed labeling pattern. These patterns are likely to be caused by the proliferation of both labeled and unlabeled terminal end bud (TEB)-resident stem cells, which deposit their progeny throughout the epithelium during ductal elongation. Labeling patterns can extend more than 8 mm in linear length and comprise many side branches, highlighting the importance of performing 3D imaging and/or macro clone analysis. Scale bar, 0.2 mm. Adapted, with permission, from Lloyd-Lewis *et al.* [24]. A schematic representation of these labeling patterns in luminal and basal clones is shown in (B). The extensive and stochastic dispersion of stem cell progeny increases the likelihood of clone convergence in studies where labeling is performed above clonal density. Clone convergence is particularly evident when using a multi-color reporter gene. In the example here (bottom panel), it is difficult to distinguish whether luminal and basal blue cells came from a single bipotent precursor or whether they arose from separate labeling events. Other technical limitations of population-based lineage-tracing approaches include periodic and promiscuous labeling by pathway-specific or pan-lineage promoters.

present in the mammary epithelium. For example, quiescent bi/multipotent MaSCs, if they exist, would not be detected by this approach [16].

Single-cell lineage-tracing has unquestionably demonstrated the immense capacity of unipotent stem cells to contribute to the development of the adult mammary epithelium, while at the same time revealing significant redundancy in the construction of each major duct [16,18] and lobuloalveolar structure [16]. Whether adult stem cells work cooperatively or competitively to achieve developmental and morphogenetic outcomes in the mammary gland is an area of active investigation and is discussed in more detail later in this review.

Multiplicity in the Mammary Gland: Roles for Potential and Quiescent Stem Cells

In addition to the cells that are responsible for the genesis and expansion of the mammary epithelium (known as professional, functional, or *bona fide* stem cells), there may also exist a population of cells in the adult breast with the capacity to behave as stem cells under particular conditions (i.e., facultative or **potential stem cells**) [1,39]. This may include (i) a subset of cells that remain quiescent during normal tissue development, and (ii) cells that are recruited under regenerative conditions [9,11,40] or in cancer [33,34]. Support for a cellular arrangement in the breast that departs from a unidirectional, top-down model is given by transplantation studies. Although it is now generally accepted that mammary repopulating cells are activated under non-homeostatic conditions [11–13], the underlying experimental observation (i.e., that not all cells are capable of repopulating the empty fat-pad [8–10]) points to the existence of a population of cells that have an intermediate or plastic nature. The physiological and pathological role of these cells, and their relationship to putative populations of quiescent MaSCs, is not immediately apparent (Figure 4). However, the notion that fate decisions within the hierarchy are not strictly unidirectional, and in some conditions can be reversed, has wide-reaching implications for oncology and regenerative medicine.

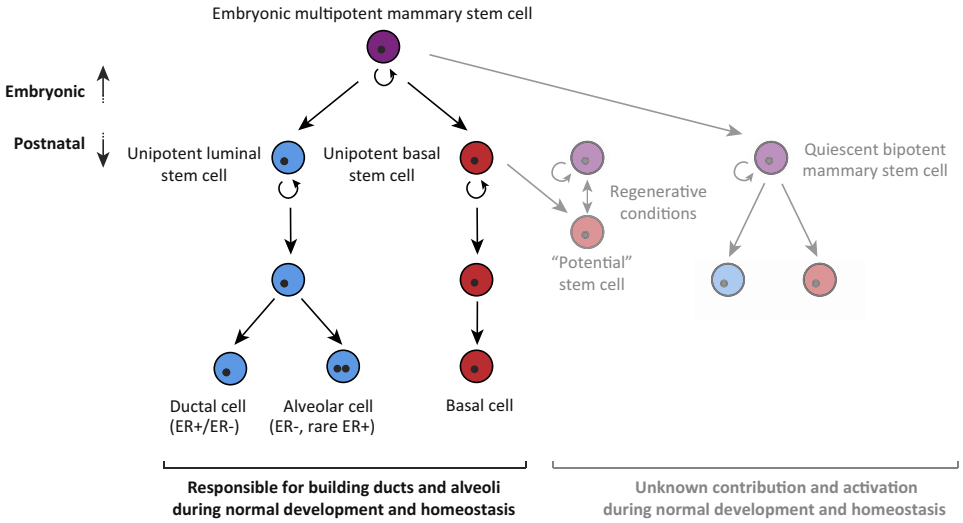


Figure 4. A Working Model of the Mammary Epithelial Cell Hierarchy. Multipotent MaSCs are present in the embryo. Although the exact stage of lineage-specification is not clear, postnatal mammary gland development (i.e., ductal and alveolar morphogenesis) is principally driven by unipotent luminal and basal mammary stem cells (MaSCs). Luminal stem cells give rise to ductal and alveolar cells that can be estrogen receptor (ER)-positive or -negative. The extent of sublineage diversity in the basal compartment, and whether there are distinct ductal and alveolar basal cells, is not yet clear. In addition to the cells responsible for building mammary ducts and alveoli under physiological conditions (left panel), various studies indicate that quiescent and potential stem cells may also reside within the adult mammary gland (right panel). Quiescent bi/multipotent MaSCs (not detected by quantitative or single-cell lineage-tracing approaches) may remain in the mammary gland after embryonic development. In addition, a plastic, intermediate cell type with properties similar to the basal cell lineage may be capable of reverting to a multipotent state under regenerative conditions. Lineage-restricted luminal and basal progenitors have also been shown to reacquire multipotency with oncogenic reprogramming. A holistic description of the cellular differentiation hierarchy in the mammary gland may need to accommodate aspects of plasticity.

A Putative Population of Quiescent MaSCs

A pool of quiescent stem cells, which have temporarily and reversibly exited the cell cycle, has been observed in several self-renewing tissues including the skin [41–43] and intestine [44]. These cells may be able to re-enter the cell cycle when required, for example upon injury [45] or homeostasis [46]. Quiescent stem cells are unlikely to be detected by conventional lineage-tracing approaches which require proliferation for clone identification [47]. As such, label-retention assays have been developed for the analysis of slow-cycling and quiescent cells [48]. DNA nucleoside analogs {e.g., bromodeoxyuridine (BrdU), 5-ethynyl-2′-deoxyuridine (EdU), and [^{3}H]-thymidine} can be used to label cells that are in cycle at the time of the pulse [47]. Alternatively, a GFP-labeled histone H2B model can be used to label specific populations of cells, with expression of H2B–GFP temporally moderated by administration of doxycycline [43,48,49]. Cells that remain labeled after a predetermined chase, known as **label-retaining cells**, are presumed to be slow-cycling/quiescent stem cells, but may also be long-lived terminally-differentiated cells [48]. Application of the H2B–GFP model to the mammary gland has identified a novel population of CD1d^{+} cells with enhanced repopulating ability upon transplantation [49]. CD1d^{+} mammary repopulating cells are also enriched for Bcl11b expression, a C2H2 zinc finger transcription factor that has independently been shown to be associated with physiological quiescence and superior repopulating activity under transplantation conditions [50]. Interestingly, neither *Cd1d* nor *Bcl11b* mRNAs are enriched in the recently-identified quiescent basal cell population defined by Lgr5 and Tspan8 expression [51]. These Lgr5^{+}Tspan8^{hi} basal cells, located within the proximal ductal tree, were also demonstrated to have enhanced repopulating activity in limiting dilution transplantation

assays [51]. Thus, these data suggest significant multiplicity, even within the putative subset of quiescent mammary repopulating cells.

Unanswered Questions: Organization, Function, and Recruitment

The proliferative demand on mammary stem and progenitor cells throughout reproductive life is substantial (Figure 1) [3,4]. Thus, the relative importance of quiescent MaSCs in normal development and homeostasis is unclear. How quiescent and potential stem cells may be recruited by specific signals in the microenvironment and their hierarchical relationship to functional stem cells is also shrouded in uncertainty. In light of the ongoing debate regarding the identity and potency of MaSCs [11–18], the fundamental necessity for cell proliferation for clone detection in lineage-tracing studies [47], and the idea that quiescent stem cells may reside at the apex of tissue hierarchies [49], one could reasonably suggest that there may be a residual population of quiescent bi/multipotent MaSCs that remain in the postnatal mammary gland after embryonic development (Figure 4). *In utero* DNA labeling has provided some support for this hypothesis, identifying long-lived label-retaining cells that are able to reversibly re-enter the cell cycle and contribute to tissue development and maintenance [46]. More recent saturation lineage-tracing, which has been able to label >95% of all cells within a single lineage, however, indicates that quiescent MaSCs (if they exist and participate in any way to tissue development and/or homeostasis), are lineage-restricted [15]. Analysis of cell division kinetics and telomere lengths in mammary epithelial populations also suggests that that each lineage is maintained by its own precursors throughout reproductive life [52].

Several important questions in this area remain unanswered. However, given the complex cellular heterogeneity in breast cancer, a long-lived and highly plastic stem cell could serve as a potential cell of origin for this disease. This highlights the importance of determining the full landscape of MaSC populations and the factors regulating their recruitment.

The MaSC Niche: An Elusive Entity or Dynamic Force?

The ability of MaSCs to rapidly and faithfully respond to developmental and homeostatic demands throughout reproductive life may be attributable to their intimate association with a specific cellular microenvironment, known as the mammary **stem cell niche**. Stem cell niches can embody discrete and highly specialized sites in particular tissues, for example the crypt base of the small intestine and the hair follicle [2]. Other tissues, including the post-pubescent mammary gland, prostate, and lung, lack an easily discernable niche, and stem cells in these organs may instead respond to more ubiquitous tissue signals [2]. In any case, reciprocal interactions between MaSCs and their mature epithelial progeny, neighboring stromal cells, and the supporting extracellular matrix undoubtedly provide the autocrine, juxtracrine, and paracrine signals that direct and adjust cell fate [19]. Extrinsic regulatory cues may include diffusible molecules (e.g., growth factors and cytokines) as well as mechanical forces (e.g., cell–cell and cell–matrix interactions) [53,54]. In this section we outline designs of mammary stem and progenitor cell distribution in the pubescent, mature, and secretory epithelium, discussing how the spatial arrangement of these cells may underpin the development and integrity of this highly dynamic tissue.

Architectural Conceptions of a MaSC Niche

The absence of a definitive molecular portrait of MaSCs, combined with uncertainties regarding their precise location within the post-pubescent mammary epithelium, has greatly impeded the analysis of prospective MaSC niches. Cell-surface signatures that facilitate the isolation of mammary repopulating cells also provide little insight into the tissue-positional cues that direct cell behavior. Early transplantation and ultrastructural studies, however, did imply that mammary repopulating cells were distributed throughout the ductal epithelium [26–28,55], positing that MaSC niches may reside in a 'suprabasal' location in the epithelial bilayer [27,56,57].

Although the precise location of stem cells within the post-pubescent breast remains unclear (Figure 1C), it is generally accepted that the TEBs of elongating ducts serve as a transient niche during puberty (Figure 1B) [5,19]. Thus, a comprehensive examination of signaling events in TEB-resident stem cells is expected to yield important insights into the pathways directing MaSC activity and fate, which may also be relevant in the post-pubescent gland. TEBs consist of an outer layer of cap cells that envelop multiple layers of inner body cells [3]. Cap and body cells are generally considered to be the precursors of mature basal and luminal epithelial lineages, respectively [19]. Cap cells have also long been hypothesized to represent an enriched population of bi/multipotent MaSCs [27,58,59]. Indeed, the alternative promoter (s-Ship) of the stem cell-associated phosphatase gene *Ship1/INPP5D*, in the mammary gland is exclusively expressed in cap cells during puberty, correlates with enhanced mammary repopulating capacity in limiting dilution transplantation assays [6]. In addition, s-Ship-expressing cap cells are strongly associated with the expression of Par3L, a protein related to the cell polarity regulator Par3 which is required for MaSC maintenance and ductal morphogenesis [60]. Recent mathematical modeling of mammary ductal elongation, however, suggests that inwardly migrating cap cells do not contribute to the luminal epithelial lineage as previously hypothesized [5]. Therefore, the precise contribution of these anatomically distinct cells to ductal morphogenesis requires further investigation. The relationship between cap cells in the TEB and unipotent MaSCs, identified by genetic lineage-tracing [11,15,16,18], is also unclear. An answer to these important questions, and a potential unifying definition of physiological MaSC potency, awaits future inducible fate-mapping studies using transgenic s-Ship and/or Par3L reporter models.

In the post-pubescent mammary gland, where TEBs have fully regressed, the location of MaSCs and their niche constituents is more ambiguous (Figure 1C). It is presumed that MaSCs, left behind by elongating TEBs during pubertal growth, are dispersed throughout the adult epithelial network, where hormonal cues stimulate further branching and the formation of alveolar-like buds and lobuloalveoli during estrous cycling and in pregnancy, respectively [61]. The notable absence of hormone receptors in mammary repopulating [62] and in MaSC-enriched basal cell populations [63] implies that paracrine interactions between hormone receptor-expressing cells and stem cells guide tissue development and homeostasis [64–68]. Multiple paracrine signaling pathways, including Wnt, EGFR, IGFR, and RANK signaling, are reported to regulate MaSC function downstream of hormone action. In addition, FGF, Hedgehog and Notch signaling pathways have also been implicated in modulating MaSC fate during different stages of mammary gland development. How the local activities of these pathways are controlled by systemic changes in hormone levels, however, remains unknown [69,70]. Nevertheless, the widespread distribution of hormone receptor-positive cells throughout the adult mammary epithelial tree [16,71] suggests that MaSCs would be able to receive and integrate these paracrine signals at most architectural locations within the ductal epithelium. Moreover, alterations in the abundance and distribution of hormone receptor-positive cells with age [71] may reflect lifetime-dependent variations in a putative MaSC niche.

MaSCs are thought to survive tissue remodeling during post-lactational involution, enabling further cycles of expansion with each subsequent pregnancy (Figure 1C–F). It is therefore tempting to speculate that MaSCs reside in the vicinity of epithelial branch points and are poised to generate the lateral branches and lobuloalveolar structures required for lactation. Fate-mapping studies using an alveolus-specific whey acidic protein (WAP)-driven Cre have also identified a population of long-lived parity-induced mammary epithelial cells (PI-MECs) that are sustained through multiple reproductive cycles [72]. These cells reside at ductal extremities in the post-parous mammary gland and contribute exclusively to the hormone receptor-negative luminal lineage in subsequent pregnancies [7,72]. Intriguingly, a recent single-cell

lineage-tracing study has revealed an unequal distribution of MaSC progeny between lobuloalveolar units in lactating mammary tissue [16]. Thus, these striking observations also support a model wherein an alveolar stem cell niche is positioned near bifurcation sites in the mature ductal epithelium. Interestingly, increased MaSC activity during pregnancy correlates with the re-expression of s-Ship specifically in basal cells at the tips of alveolar buds, suggesting the emergence of a transient stem cell niche during lobuloalveologenesis [6].

MaSC Niche Dynamics

As described earlier in this review, distinct adult MaSCs are postulated to fulfill the proliferative and homeostatic demands of the mammary gland (Figure 4) [19]. The degree to which the heterogeneity in the MaSC compartment is intrinsic or a result of microenvironmental cues, however, is not known. A recent single-cell lineage-tracing study, which employed quantitative volumetric analysis to determine the contribution of a single labeled MaSC to ductal morphogenesis, estimated that at least 35 lineage-restricted MaSCs actively and stochastically contribute to the development of each major duct during puberty [16]. A subsequent study, also using quantitative lineage-tracing at clonal density, put this number at 260 lineage-restricted MaSCs per TEB, leading to the suggestion that most TEB cells can function as lineage-committed MaSCs [18]. Discrepancies between these two studies may reflect differing functional definitions of MaSCs as well as the quantitative and mathematical platforms and assumptions for analysis. Quantitative lineage-tracing studies also suggest that molecularly heterogeneous populations of TEB-resident MaSCs function as single equipotent pools, colonizing ductal branches through stochastic neutral drift dynamics [18]. Random segregation during successive rounds of TEB bifurcation mediates the unequal distribution of MaSC progeny between adjacent ductal structures, leading to clonal enrichment or extinction over time [18], supporting previous observations of clonal labeling patterns [16]. Furthermore, single-cell lineage-tracing has shown that most lactational alveoli comprise the progeny of more than a single unipotent MaSC, indicating that a pool of lineage-restricted alveolar MaSCs also contribute to alveolar morphogenesis during pregnancy and lactation [16]. These early applications of quantitative and single-cell lineage-tracing approaches in the mammary gland [16,18] have provided unprecedented insights into clonal dynamics and stem/progenitor heterogeneity and multiplicity, heralding a new era in our investigation and understanding of normal and malignant stem cells in the breast.

Concluding Remarks

In this review we have examined the properties of potency, dynamics, and plasticity in adult MaSCs and the respective technologies that have underpinned key experimental observations. Although this area has received considerable attention over the past decade, many questions remain unanswered (see Outstanding Questions). At the center of this enquiry is whether MaSCs in the adult breast are unipotent, bipotent, or something less discordant.

Stem cells are defined by their functional abilities, that is: proliferation, self-maintenance, the production of a large number of differentiated progeny, tissue regeneration/repair, and flexibility within these states [39]. The challenge thus far has been how to study the functionality of a cell without inadvertently altering its function. Lineage tracing has come a long way in this respect [11–16,18]. The refinement of lineage-tracing approaches and the application of other novel experimental models and methods for marking, visualizing, and profiling individual cells (Box 1) will continue to provide important insights in this field. The question then becomes, what level of evidence is required to achieve a consensus?

Outstanding Questions

Is there a distinct and universal molecular signature for adult MaSCs? Or do stem cells reside along a wide and dynamic spectrum of activities?

How do lineage-restricted cells reacquire a multipotent fate under regenerative conditions or in particular cancers?

Is there a hierarchical relationship between functional and putative quiescent mammary stem cell populations?

Do MaSCs reside within a stem cell niche in the post-pubescent mammary gland? Is their relationship within the niche cooperative or competitive?

Do different niche environments exist for quiescent and functional MaSC populations, similar to those observed in the intestine and hair follicle?

Do polyclonal ducts and alveoli drift towards monoclonality with age?

What level of evidence will be required to achieve a unifying definition of physiological MaSC potency?

Acknowledgments

This work was supported by the Medical Research Council (MR/J001023/1 to C.J.W. and B.L-L.), the Wellcome Trust (105377/Z/14/Z to O.B.H.), the National Health and Medical Research Council (1071074 to F.M.D.), and a University of

Queensland Early Career Researcher Grant (UQECR1718865 to F.M.D.). We thank Dr Silvia Fre for her valuable feedback on this manuscript, and apologize to all investigators whose work could not be cited owing to space limitations.

References

1. Visvader, J.E. and Clevers, H. (2016) Tissue-specific designs of stem cell hierarchies. *Nat. Cell Biol.* 18, 349–355
2. O'Brien, L.E. and Bilder, D. (2013) Beyond the niche: tissue-level coordination of stem cell dynamics. *Annu. Rev. Cell Dev. Biol.* 29, 107–136
3. Macias, H. and Hinck, L. (2012) Mammary gland development. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 533–557
4. Watson, C.J. and Khaled, W.T. (2008) Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development* 135, 995–1003
5. Paine, I. *et al.* (2016) A geometrically-constrained mathematical model of mammary gland ductal elongation reveals novel cellular dynamics within the terminal end bud. *PLoS Comput. Biol.* 12, e1004839
6. Bai, L. and Rohrschneider, L.R. (2010) s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes Dev.* 24, 1882–1892
7. Chang, T.H.-T. *et al.* (2014) New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. *Breast Cancer Res.* 16, R1
8. Stingl, J. *et al.* (2006) Purification and unique properties of mammary epithelial stem cells. *Nature* 439, 993–997
9. Shackleton, M. *et al.* (2006) Generation of a functional mammary gland from a single stem cell. *Nature* 439, 84–88
10. Sleeman, K.E. *et al.* (2006) CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast Cancer Res.* 8, R7
11. Van Keymeulen, A. (2011) Distinct stem cells contribute to mammary gland development and maintenance. *Nature* 479, 189–193
12. de Visser, K.E. *et al.* (2012) Developmental stage-specific contribution of LGR5$^+$ cells to basal and luminal epithelial lineages in the postnatal mammary gland. *J. Pathol.* 228, 300–309
13. van Amerongen, R. (2012) Developmental stage and time dictate the fate of Wnt/β-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* 11, 387–400
14. Rios, A.C. *et al.* (2014) *In situ* identification of bipotent stem cells in the mammary gland. *Nature* 506, 322–327
15. Wuidart, A. *et al.* (2016) Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes Dev.* 30, 1261–1277
16. Davis, F.M. *et al.* (2016) Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat. Commun.* 7, 13053
17. Wang, D. *et al.* (2015) Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* 517, 81–84
18. Scheele, C.L.G.J. *et al.* (2017) Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* 542, 313–317
19. Sreekumar, A. *et al.* (2015) The mammary stem cell hierarchy: a looking glass into heterogeneous breast cancer landscapes. *Endocr. Relat. Cancer* 22, T161–T176
20. Visvader, J.E. and Stingl, J. (2014) Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.* 28, 1143–1158
21. Nassar, D. and Blanpain, C. (2016) Cancer stem cells: basic concepts and therapeutic implications. *Annu. Rev. Pathol. Mech. Dis.* 11, 47–76
22. DeOme, K.B. *et al.* (1959) Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer Res.* 19, 515–520
23. Rios, A.C. *et al.* (2016) Essential role for a novel population of binucleated mammary epithelial cells in lactation. *Nat. Commun.* 7, 11400
24. Lloyd-Lewis, B. *et al.* (2016) Imaging the mammary gland and mammary tumours in 3D: optical tissue clearing and immunofluorescence methods. *Breast Cancer Res.* 18, 127
25. Sale, S. and Pavelic, K. (2015) Mammary lineage tracing: the coming of age. *Cell. Mol. Life Sci.* 72, 1577–1583
26. Hoshino, K. (1964) Regeneration and growth of quantitatively transplanted mammary glands of normal female mice. *Anat. Rec.* 150, 221–236
27. Smith, G.H. and Medina, D. (1988) A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. *J. Cell Sci.* 90, 173–183
28. Daniel, C. *et al.* (1968) The *in vivo* life span of normal and preneoplastic mouse mammary glands: a serial transplantation study. *Proc. Natl. Acad. Sci. U. S. A.* 61, 53–60
29. Kordon, E.C. and Smith, G.H. (1998) An entire functional mammary gland may comprise the progeny from a single cell. *Development* 125, 1921–1930
30. Ehmann, U.K. *et al.* (1987) Cultured mouse mammary epithelial cells: normal phenotype after implantation. *J. Natl. Cancer Inst.* 78, 751–757
31. Smith, G.H. (1996) Experimental mammary epithelial morphogenesis in an *in vivo* model: Evidence for distinct cellular progenitors of the ductal and lobular phenotype. *Breast Cancer Res. Treat.* 39, 21–31
32. Prater, M.D. *et al.* (2014) Mammary stem cells have myoepithelial cell properties. *Nat. Cell Biol.* 16, 942–950
33. Van Keymeulen, A. (2015) Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. *Nature* 525, 119–123
34. Koren, S. *et al.* (2015) PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours. *Nature* 525, 114–118
35. Rios, A.C. *et al.* (2016) The complexities and caveats of lineage tracing in the mammary gland. *Breast Cancer Res.* 18, 116
36. Rodilla, V. *et al.* (2015) Luminal progenitors restrict their lineage potential during mammary gland development. *PLoS Biol.* 13, e1002069
37. Lafkas, D. *et al.* (2013) Notch3 marks clonogenic mammary luminal progenitor cells *in vivo*. *J. Cell Biol.* 203, 47–56
38. Kozar, S. *et al.* (2013) Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell* 13, 626–633
39. Potten, C.S. and Loeffler, M. (1990) Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* 110, 1001–1020
40. Aloia, L. *et al.* (2016) Cellular plasticity in the adult liver and stomach. *J. Physiol.* 594, 4815–4825
41. Cotsarelis, G. *et al.* (1990) Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* 61, 1329–1237
42. Horsley, V. *et al.* (2008) NFATc1 balances quiescence and proliferation of skin stem cells. *Cell* 132, 299–310
43. Tumbar, T. *et al.* (2004) Defining the epithelial stem cell niche in skin. *Science* 303, 359–363
44. Buczacki, S.J.A. *et al.* (2013) Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* 495, 65–69
45. Ito, M. *et al.* (2005) Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. *Nat. Med.* 11, 1351–1354
46. Boras-Granic, K. *et al.* (2014) Embryonic cells contribute directly to the quiescent stem cell population in the adult mouse mammary gland. *Breast Cancer Res.* 16, 487
47. Li, L. and Clevers, H. (2010) Coexistence of quiescent and active adult stem cells in mammals. *Science* 327, 542–545
48. Fuchs, E. and Horsley, V. (2011) Ferreting out stem cells from their niches. *Nat. Cell Biol.* 13, 513–518
49. dos Santos, C.O. *et al.* (2013) Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 110, 7123–7130

50. Cai, S. *et al.* (2017) A quiescent Bcl11b high stem cell population is required for maintenance of the mammary gland. *Cell Stem Cell* 20, 247–260
51. Fu, N. *et al.* (2017) Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive. *Nat. Cell Biol.* 19, 164–176
52. Giraddi, R.R. *et al.* (2015) Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. *Nat. Commun.* 6, 8487
53. Inman, J.L. *et al.* (2015) Mammary gland development: cell fate specification, stem cells and the microenvironment. *Development* 142, 1028–1042
54. Howard, B.A. and Lu, P. (2014) Stromal regulation of embryonic and postnatal mammary epithelial development and differentiation. *Semin. Cell Dev. Biol.* 25/26, 43–51
55. Young, L.J.T. *et al.* (1971) The influence of host and tissue age on life span and growth rate of serially transplanted mouse mammary gland. *Exp. Gerontol.* 6, 49–56
56. Chepko, G. and Smith, G.H. (1997) Three division-competent, structurally-distinct cell populations contribute to murine mammary epithelial renewal. *Tissue Cell* 29, 239–253
57. Chepko, G. and Dickson, R.B. (2003) Ultrastructure of the putative stem cell niche in rat mammary epithelium. *Tissue Cell* 35, 83–93
58. Williams, J.M. and Daniel, C.W. (1983) Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. *Dev. Biol.* 97, 274–290
59. Srinivasan, K. *et al.* (2003) Netrin-1/neogenin interaction stabilizes multipotent progenitor cap cells during mammary gland morphogenesis. *Dev. Cell* 4, 371–382
60. Huo, Y. and Macara, I.G. (2014) The Par3-like polarity protein Par3L is essential for mammary stem cell maintenance. *Nat. Cell Biol.* 16, 529–537
61. Brisken, C. and O'Malley, B. (2010) Hormone action in the mammary gland. *Cold Spring Harb. Perspect. Biol.* 2, a003178
62. Sleeman, K.E. *et al.* (2007) Dissociation of estrogen receptor expression and *in vivo* stem cell activity in the mammary gland. *J. Cell Biol.* 176, 19–26
63. Asselin-Labat, M.-L. (2006) Steroid hormone receptor status of mouse mammary stem cells. *J. Natl. Cancer Inst.* 98, 1011–1014
64. Beleut, M. *et al.* (2010) Two distinct mechanisms underlie progesterone-induced proliferation in the mammary gland. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2989–2894
65. Asselin-Labat, M.-L. (2010) Control of mammary stem cell function by steroid hormone signalling. *Nature* 465, 798–802
66. Cai, C. *et al.* (2014) R-spondin1 is a novel hormone mediator for mammary stem cell self-renewal. *Genes Dev.* 28, 2205–2218
67. Rajaram, R.D. *et al.* (2015) Progesterone and Wnt4 control mammary stem cells via myoepithelial crosstalk. *EMBO J.* 34, 641–652
68. Joshi, P. a *et al.* (2010) Progesterone induces adult mammary stem cell expansion. *Nature* 465, 803–807
69. Rosen, J.M. and Roarty, K. (2014) Paracrine signaling in mammary gland development: what can we learn about intratumoral heterogeneity? *Breast Cancer Res.* 16, 202
70. Brisken, C. and Ataca, D. (2015) Endocrine hormones and local signals during the development of the mouse mammary gland. *Wiley Interdiscip. Rev. Dev. Biol.* 4, 181–195
71. Ismail, P.M. *et al.* (2002) A novel LacZ reporter mouse reveals complex regulation of the progesterone receptor promoter during mammary gland development. *Mol. Endocrinol.* 16, 2475–2489
72. Wagner, K.-U. *et al.* (2002) An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. *Development* 129, 1377–1386
73. Tainaka, K. *et al.* (2014) Whole-body imaging with single-cell resolution by tissue decolorization. *Cell* 159, 911–924
74. Ke, M.-T. *et al.* (2013) SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nat. Neurosci.* 16, 1154–1161
75. Davis, F.M. *et al.* (2015) Essential role of Orai1 store-operated calcium channels in lactation. *Proc. Natl. Acad. Sci. U. S. A.* 112, 5827–5832
76. Ritsma, L. *et al.* (2013) Surgical implantation of an abdominal imaging window for intravital microscopy. *Nat. Protoc.* 8, 583–594
77. Zomer, A. *et al.* (2013) Intravital imaging of cancer stem cell plasticity in mammary tumors. *Stem Cells* 31, 602–606
78. Kumar, P. *et al.* (2017) Understanding development and stem cells using single cell-based analyses of gene expression. *Development* 144, 17–32
79. McKenna, A. *et al.* (2016) Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* 353, aaf7907
80. Perli, S.S.D. *et al.* (2016) Continuous genetic recording with self-targeting CRISPR-Cas in human cells. *Science* 511, 53058
81. Clevers, H. (2016) Modeling development and disease with organoids. *Cell* 165, 1586–1597
82. Jardé, T. *et al.* (2016) Wnt and neuregulin1/ErbB signalling extends 3D culture of hormone responsive mammary organoids. *Nat. Commun.* 7, 13207
83. Cardiff, R.D. and Wellings, S.R. (1999) The comparative pathology of human and mouse mammary glands. *J. Mammary Gland Biol. Neoplasia* 4, 105–122
84. Oakes, S.R. *et al.* (2014) The mammary cellular hierarchy and breast cancer. *Cell. Mol. Life Sci.* 71, 4301–4324

© 2018. Published by The Company of Biologists Ltd | Development (2018) 145, dev164079. doi:10.1242/dev.164079

STEM CELLS AND REGENERATION

RESEARCH REPORT

Neutral lineage tracing of proliferative embryonic and adult mammary stem/progenitor cells

Bethan Lloyd-Lewis[1,*,§,¶], Felicity M. Davis[1,‡,§,¶], Olivia B. Harris[1,2], Jessica R. Hitchcock[1] and Christine J. Watson[1,2,¶]

ABSTRACT

Mammary gland development occurs over multiple phases, beginning in the mammalian embryo and continuing throughout reproductive life. The remarkable morphogenetic capacity of the mammary gland at each stage of development is attributed to the activities of distinct populations of mammary stem cells (MaSCs) and progenitor cells. However, the relationship between embryonic and adult MaSCs, and their fate during different waves of mammary gland morphogenesis, remains unclear. By employing a neutral, low-density genetic labelling strategy, we characterised the contribution of proliferative stem/progenitor cells to embryonic, pubertal and reproductive mammary gland development. Our findings further support a model of lineage restriction of MaSCs in the postnatal mammary gland, and highlight extensive redundancy and heterogeneity within the adult stem/progenitor cell pool. Furthermore, our data suggest extensive multiplicity in their foetal precursors that give rise to the primordial mammary epithelium before birth. In addition, using a single-cell labelling approach, we revealed the extraordinary capacity of a single embryonic MaSC to contribute to postnatal ductal development. Together, these findings provide tantalising new insights into the disparate and stage-specific contribution of distinct stem/progenitor cells to mammary gland development.

KEY WORDS: Mammary gland development, Embryonic mammary stem cells, Adult mammary stem cells, Lineage tracing

INTRODUCTION

Mammary gland development is a complex and multi-stage process that begins in the embryo and continues throughout the reproductive life of female mammals (Cowin and Wysolmerski, 2010; Gjorevski and Nelson, 2011; Hinck and Silberstein, 2005). This process commences with the formation of two milk lines from the overlying ectoderm on embryonic day (E) 10.5 in mice, and the asynchronous appearance of five pairs of placodes at specific and symmetric locations between the fore- and hindlimbs by E11.5 (Hens and Wysolmerski, 2005). These ectodermal placodes develop asynchronously and invaginate to form the mammary bud by E13.5, followed by the formation of an epithelial tubular sprout by E15.5-16.5 that invades the underlying mammary fat pad precursor. Contact with the developing fat pad initiates a phase of branching morphogenesis, resulting in the formation of the primordial ductal tree by E18.5, prior to birth (Veltmaat, 2017; Veltmaat et al., 2003; Watson and Khaled, 2008).

In the weeks immediately after birth, growth of the ductal tree is commensurate with body growth and it is not until puberty that ductal structures begin to elongate rapidly and invade the empty fat pad, driven by hormonal and growth factor signalling in the micro-environment (Hinck and Silberstein, 2005). This process, known as ductal morphogenesis, is orchestrated by proliferation of adult mammary stem and progenitor cells within the distal terminal end bud (TEB) structures (Bai and Rohrschneider, 2010; Paine et al., 2016; Sreekumar et al., 2015). After pubertal growth is complete, the mammary epithelium re-enters a phase of balanced proliferation, with only minor growth and remodelling occurring with cyclical ovarian hormone stimulation. However, rapid expansion of the epithelium again occurs during pregnancy and lactation, when adult MaSCs proliferate to form lobuloalveolar structures capable of producing and expelling milk for neonatal nourishment (Davis et al., 2016; Lloyd-Lewis et al., 2017; Sreekumar et al., 2015).

Despite their essential role in pre- and postnatal mammary gland development, studies to determine the molecular identity and differentiation potential of MaSCs have yielded conflicting results (for recent reviews, see Sreekumar et al., 2015; Lloyd-Lewis et al., 2017). Recently, both saturation and single-cell genetic lineage-tracing studies have demonstrated that lineage-restricted MaSCs appear to drive postnatal mammary development under physiological conditions (Davis et al., 2016; Scheele et al., 2017; Wuidart et al., 2016). However, these studies also demonstrated significant redundancy and heterogeneity within the adult MaSC compartment, and the differential and stage-specific contribution of diverse stem/progenitor cells in the breast is still emerging (Bach et al., 2017; Cai et al., 2017; Van Keymeulen et al., 2017; Wang et al., 2017). In this study, we employed a low-density, neutral, genetic labelling strategy to further investigate the extent and nature of the contribution of proliferative stem/progenitor cells to embryonic, pubertal and reproductive mammary development.

RESULTS AND DISCUSSION

A pool of lineage-biased adult stem/progenitor cells propel ductal elongation during puberty

Recently, genetic lineage-tracing studies in the mouse mammary gland have achieved *in vivo* indelible marking of specific populations of cells (characterised by their expression of nominated genes at specific developmental stages) and the

[1]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK. [2]Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 1QR, UK.
*Present address: Department of Genetics and Developmental Biology, Institut Curie, 75248 PARIS Cedex 05, France. ‡Present address: Mater Research Institute, The University of Queensland Translational Research Institute, Queensland 4102, Australia. §These authors contributed equally to this work.

¶Authors for correspondence (bethan.lloyd-lewis@curie.fr; f.davis@uq.edu.au; cjw53@cam.ac.uk)

B.L., 0000-0001-6511-1818; F.M.D., 0000-0001-9112-118X; O.B.H., 0000-0002-2641-6472; C.J.W., 0000-0002-8548-5902

Received 3 February 2018; Accepted 19 June 2018

Development (2018) 145, dev164079. doi:10.1242/dev.164079

subsequent analysis of the progeny of proliferative labelled cells after an appropriate chase (Sale and Pavelic, 2015). Targeted cell populations include those temporally or stably expressing: keratin (K) 5 (Rios et al., 2014; Van Keymeulen et al., 2011), K14 (Rios et al., 2014; Tao et al., 2014; Van Keymeulen et al., 2011; Wuidart et al., 2016), K8 (Tao et al., 2014; Van Keymeulen et al., 2011; Wuidart et al., 2016), K18 (Van Keymeulen et al., 2011), K19 (Wuidart et al., 2016), Elf5 (Rios et al., 2014), Lgr5 (de Visser et al., 2012; Fu et al., 2017; Rios et al., 2014; Van Keymeulen et al., 2011; Wuidart et al., 2016), Lgr6 (Blaas et al., 2016; Wuidart et al., 2016), Sox9 (Wang et al., 2017; Wuidart et al., 2016), Axin2 (van Amerongen et al., 2012), Notch1 (Rodilla et al., 2015), Notch2 (Šale et al., 2013), Notch3 (Lafkas et al., 2013), WAP (Chang et al., 2014), Acta2 (Prater et al., 2014), p63 (Sreekumar et al., 2017), Procr (Wang et al., 2015), prominin 1 (Wang et al., 2017) and ER (Van Keymeulen et al., 2017). However, although providing valuable information on mammary development and the epithelial differentiation hierarchy, these models have relied on prior assumptions regarding the specificity and consistency of the expression of the chosen gene promoters, and have generated conflicting results.

In this study, we have employed a neutral genetic labelling strategy for lineage analysis in the mammary gland using $R26^{CreERT2}$;$R26^{Confetti}$ mice (Fig. 1A) (Davis et al., 2016; Li et al., 2016; Scheele et al., 2017). Administration of a low dose of tamoxifen induces the stochastic expression of up to four fluorescent proteins (FPs) (Fig. 1A). Importantly, FP expression can occur in any cell, overcoming issues pertaining to the requisite high-level Cre specificity inherent to other models (discussed by Wuidart et al., 2016; Davis et al., 2016 ; Lloyd-Lewis et al., 2017).

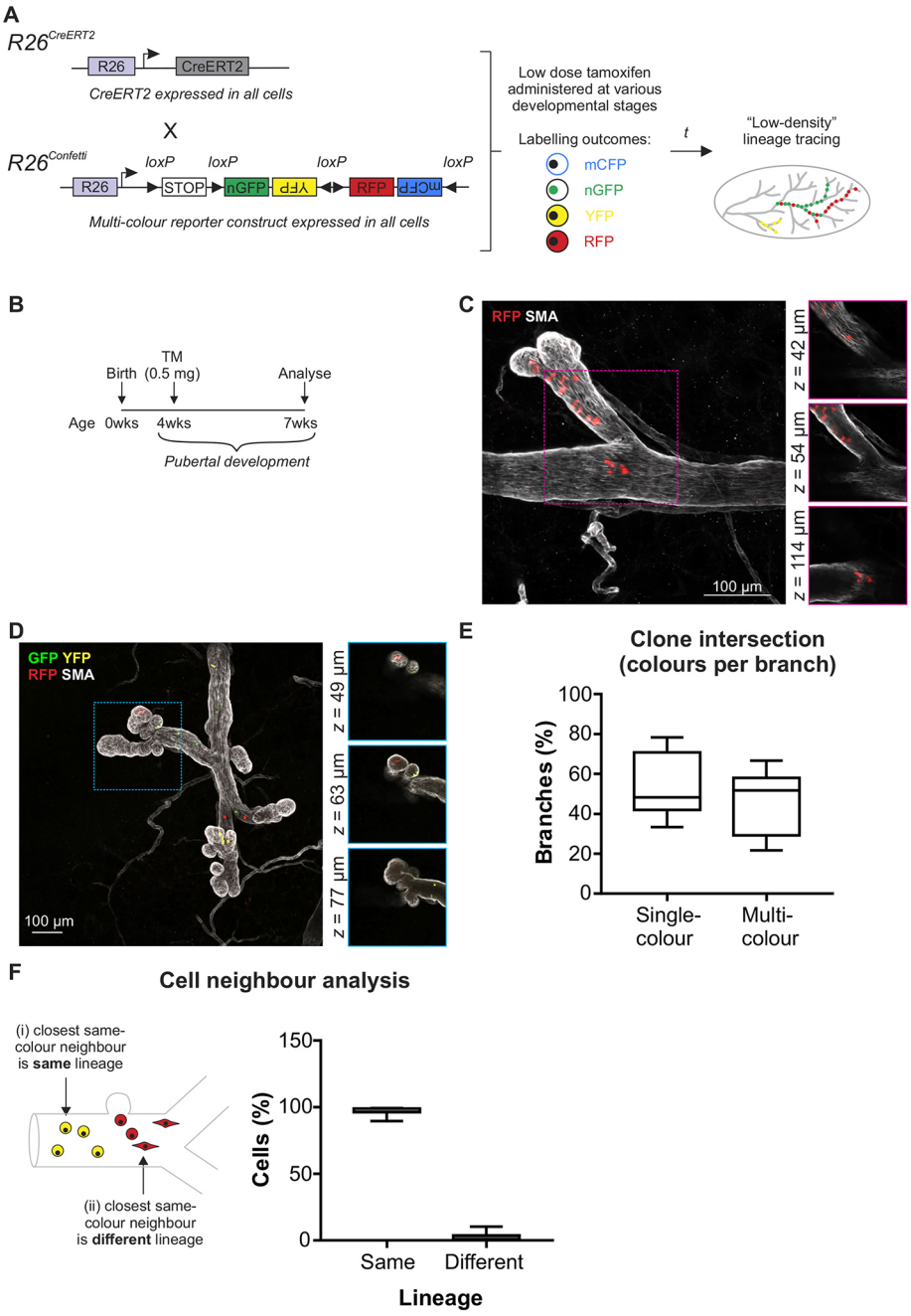


Fig. 1. Lineage tracing during branching morphogenesis. (A) The $R26^{CreERT2}$;$R26^{Confetti}$ model. $R26^{CreERT2}$ mice (expressing inducible Cre-recombinase in all cells) were crossed to $R26^{Confetti}$ mice (expressing a conditional multicolour reporter in all cells) to generate double hemizygous mice. Administration of low-dose tamoxifen produced stochastic genetic labelling of cells at relatively low density. Labelling outcomes include membranous CFP (mCFP), nuclear GFP (nGFP), cytosolic YFP (YFP) or cytosolic RFP (RFP); however, CFP^+ clones (Fig. S2) were under-represented (Davis et al., 2016) and were not analysed. (B) For lineage tracing during branching morphogenesis, tamoxifen was administered (4 weeks) and tissue harvested (7 weeks). (C,D) Example of single-colour branches (C) and multicoloured branches (D). Images show maximum-intensity *z*-projections and optical slices of a region of interest (boxed and enlarged in the right-hand panels). (E) The percentage of single- and multicolour branches in pubertal mice. (F) Cell neighbour analysis revealed that the majority of FP^+ cells had a same-colour FP^+ neighbour that was the same lineage. Box extends from the 25th to 75th percentiles and whiskers indicate minimum to maximum values. Data are from 1419 cells distributed across 130 branches from randomly selected 3D images (five mice).

Development (2018) 145, dev164079. doi:10.1242/dev.164079

Neutral labelling of proliferative cells at clonal density (where the chance of clone convergence is extremely low) has previously been described using the $R26^{CreERT2};R26^{Confetti}$ model (using an 'ultra-low' dose of tamoxifen; 0.2 mg per 25 g body weight) (Scheele et al., 2017) and the $R26^{[CAJ30}$ model (Davis et al., 2016). Using these models combined with 3D imaging, all of the progeny of a single labelled cell can be analysed with confidence. These studies revealed that lineage-restricted stem/progenitor cells orchestrate ductal (Davis et al., 2016; Scheele et al., 2017) and alveolar (Davis et al., 2016) mammary morphogenesis. However, they also revealed extraordinary multiplicity in the MaSC compartment and thus their power to capture the full spectrum of mammary stem/progenitor cells is limited.

In the current study, we injected pubertal $R26^{CreERT2};R26^{Confetti}$ mice with 0.5 mg tamoxifen (~35 µg/g) to achieve low-density labelling in the mammary epithelium (Fig. 1B and Fig. S1A). This dose is approximately fourfold higher than previous studies using 'ultra-low' tamoxifen dosing in puberty (Scheele et al., 2017). Using this approach, we observed mammary branches that contained labelled cells of a single colour (Fig. 1C) as well as branches comprising two or more colours (Fig. 1D), as expected. No labelling was observed in control vehicle-injected mice (Fig. S1B). Quantification of the number of single- and multicoloured branches indicated that, under these conditions, the likelihood of clone convergence is at least 50% (Fig. 1E); this number may be even higher, as distinct coincident labelling events of the same colour cannot be distinguished.

Consistent with previous reports (Davis et al., 2016), we observed stochastic dispersion of labelled cell progeny throughout the developing ducts (Fig. 1C,D and Fig. S3). This labelling pattern is likely to have arisen from the deposition of labelled progeny along developing ducts by proliferative labelled cells in elongating TEBs (Bai and Rohrschneider, 2010; Davis et al., 2016). As ductal elongation and side branching occur as the result of cell proliferation by stem/progenitor cells within both TEB and ductal structures and the admixing of clonal progeny (Fu et al., 2017; Rios et al., 2014; Wang et al., 2015), we employed a cell-neighbour analysis to assess lineage potential (Fig. 1F; see supplementary Materials and Methods). A striking majority of same-colour cell neighbours consisted of cells of the same lineage, providing further evidence of physiological lineage bias in the postnatal mammary gland (Davis et al., 2016; Scheele et al., 2017; Van Keymeulen et al., 2011; Wang et al., 2017; Wuidart et al., 2016).

Alveolar morphogenesis is driven by a pool of lineage-biased adult stem/progenitor cells

Low-density labelling using the $R26^{CreERT2};R26^{Confetti}$ model was also used for lineage analysis during alveolar morphogenesis (Fig. 2, Figs S4 and S5). $R26^{CreERT2};R26^{Confetti}$ mice were injected with low-dose tamoxifen (1 mg per mouse; ~40-50 µg/g) (Fig. S5A), mated and tissue harvested during lactation (Fig. 2A). Under these conditions, a large number of single-colour alveoli were observed (Fig. 2B,C and Fig. S4) with fewer multicoloured alveoli (Fig. 2D). No FP^+ cells were observed in control mice (Fig. S5B). Analysis of individual alveolar units revealed the vast majority (96.6%) of alveoli were single-coloured (Fig. 2E). Of the single-coloured alveoli, only 0.1% contained both luminal and basal cells of the same colour (Fig. 2F). Thus, these data support previous lineage-tracing studies using a different neutral model at single cell density showing lineage restriction during alveolar morphogenesis (Davis et al., 2016). Previous single cell lineage-tracing studies, which quantified only very large labelled clones (containing hundreds of labelled cells, with each clone presumably arising from a single MaSC), demonstrated that most alveoli comprise the progeny of a pool of lineage-restricted cells (Davis et al., 2016). Analysis of the number of partially versus fully populated alveoli in this low-density model revealed a seemingly higher rate of polyclonality (Fig. 2G). This is likely due to the inclusion of small, medium and large clones in the current study, representing the wider spectrum of stem and progenitor cell divisions.

Neutral lineage tracing supports the presence of lineage-biased embryonic stem/progenitor cells by late embryogenesis

Although MaSCs appear to be lineage restricted postnatally, numerous studies have suggested that their foetal precursors are multipotent (Boras-Granic et al., 2014; Fu et al., 2017; Rodilla et al., 2015; Spike et al., 2012; Trejo et al., 2017; Van Keymeulen et al., 2011; Wang et al., 2015). Embryonic MaSCs display multipotential activity in *in vitro* and transplantation assays, and increase dramatically in number during this developmental window (Spike et al., 2012). Moreover, *in vivo* population-based fate mapping has shown that all mammary epithelial lineages derive from embryonic K14-expressing stem/progenitor cells labelled at E17 (Van Keymeulen et al., 2011), an observation reinforced by other lineage-tracing studies using different gene promoters (Fu et al., 2017; Rodilla et al., 2015; Trejo et al., 2017; Wang et al., 2015). However, whether embryonic MaSCs are truly multipotent or whether they consist of populations of distinct progenitors that are already committed to give rise to different lineages after birth, has remained subject to deliberation. In addition, it has remained unclear when putative multipotent embryonic MaSCs become lineage restricted. Recent studies using low-density lineage tracing of embryonic cells have shed new light on this debate (Lilja et al., 2018; Wuidart et al., 2018). Clonal labelling of K14-expressing cells at E13 (when K14 appears to be universally expressed in luminal and basal lineages) points to the existence of multipotent stem cells at this stage of mammogenesis (Wuidart et al., 2018). However, at birth, segregation of basal and luminal lineages appears to be complete, with K5-expressing (Wuidart et al., 2018) and Acta2-expressing (Lilja et al., 2018) cells exclusively giving rise to basal progeny, and Notch1-expressing (Lilja et al., 2018) cells exclusively giving rise to luminal progeny. As targeted promoters may be differentially expressed in the neonatal and prenatal mammary gland (Sun et al., 2010; Boras-Granic et al., 2014; Trejo et al., 2017), definitive determination of the potential of embryonic MaSCs and their perinatal lineage segregation requires a neutral and inducible approach to labelling that is independent of these promoters. In addition, it has been demonstrated (Rios et al., 2016) that enzymatic digestion prior to 3D visualisation (Lilja et al., 2018; Wuidart et al., 2018) can alter tissue architecture and cell morphology, potentially confounding lineage-tracing outcomes. Unequivocal lineage determination must therefore include studies that employ methods of 3D visualisation that are void of proteolytic digestion (Lloyd-Lewis et al., 2016). To address this, a single low-dose of tamoxifen (33 µg per g maternal body weight) was administered by oral gavage to pregnant transgenic mice to induce incontrovertible neutral labelling in $R26^{CreERT2};R26^{Confetti}$ embryos at E16.5-E17.5. This route of delivery, although subject to first-pass metabolism, is reported to have less embryonic toxicity and more uniform recombination by Cre (Park et al., 2008). Mammary glands of offspring labelled *in utero* were subsequently visualised in 3D (without prior proteolytic digestion) (Fig. 3A). Using this approach, we observed large regions of labelled cells, some spanning from the

Development (2018) 145, dev164079. doi:10.1242/dev.164079

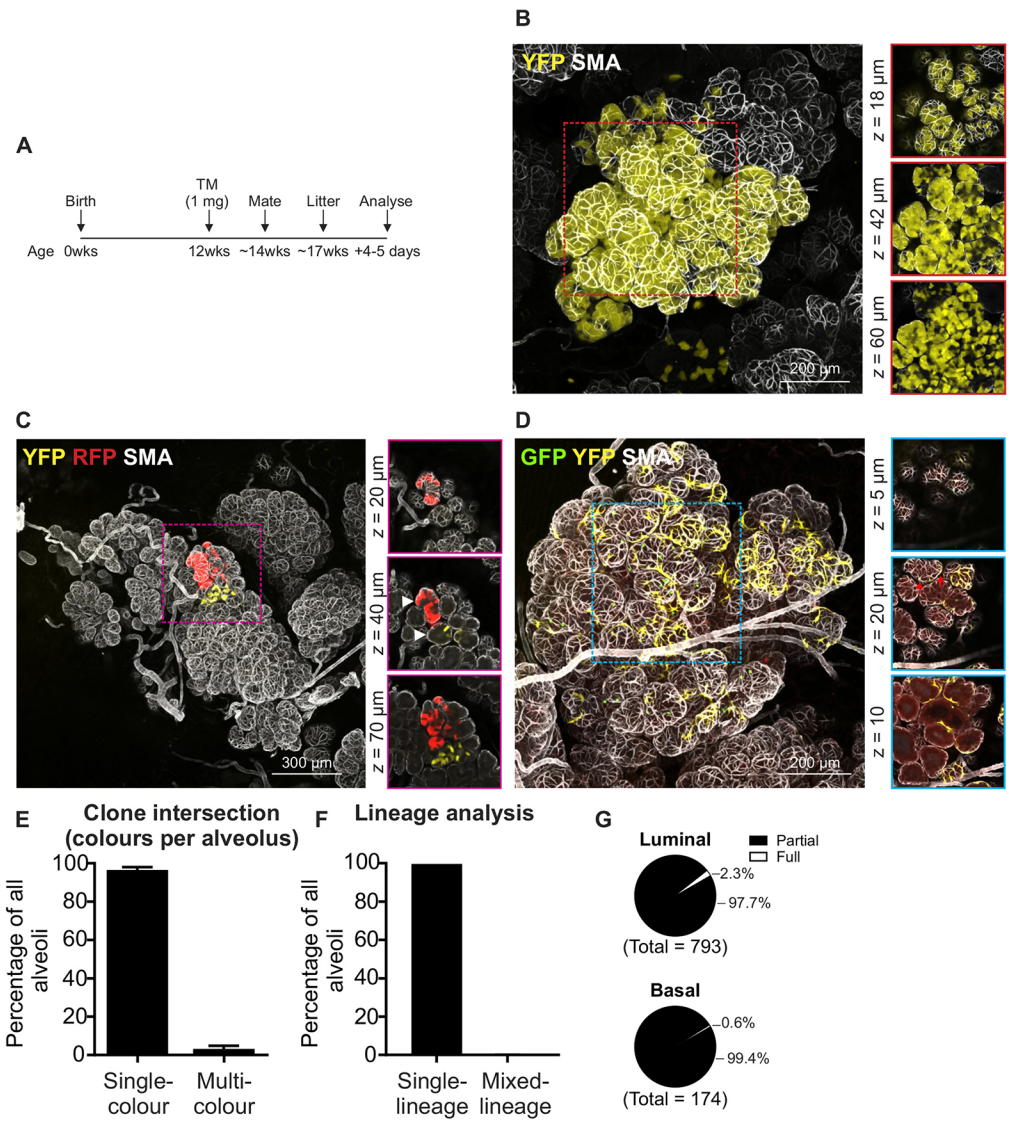


Fig. 2. Lineage tracing after alveolar morphogenesis. (A) Tamoxifen was administered (~12 weeks), mice were mated (~14 weeks) and tissue harvested during lactation. (B-D) Example of single-colour luminal YFP$^+$ alveoli (B), single-colour luminal YFP$^+$ and RFP$^+$ adjacent alveoli within a larger lobuloalveolar structure (arrowheads indicate different alveoli) (C) and multicolour basal GFP$^+$ and YFP$^+$ alveoli (D; red arrows show GFP$^+$ and YFP$^+$ cells within a single alveolus). Images show maximum-intensity *z*-projections and optical slices of a region of interest (boxed and enlarged in the right-hand panels). (E) Graph (data are mean±s.e.m.) showing the percentage of single- and multicoloured alveoli; a lower rate of clone convergence is observed in this model following expansion during gestation and lactation. (F) Graph (data are mean±s.e.m.) showing the percentage of FP$^+$ alveoli in which the same-colour cells were the same lineage (i.e. all luminal or all basal) or where same-colour cells were mixed lineage (both luminal and basal). (G) Fraction of alveoli that were fully populated by single-colour FP$^+$ cells of a single lineage (full) versus those populated by both single- or multicoloured FP$^+$ cells and/or unlabelled cells of a single lineage (partial). Data represent 1016 alveoli from randomly selected 3D images (three mice).

nipple to the outer reaches of the fad pad (Fig. 3B-F, and Fig. S6). 3D imaging of areas proximal to the nipple revealed that these regions always comprised cells of two or more colours (Fig. 3C). Ductal branches in distal regions, however, more commonly contained cells of a single colour (Fig. 3D-F versus Fig. S7). Quantification of the number of single- and multicoloured branches in nipple and distal regions confirmed this observation, and indicated that the likelihood of clone convergence under these conditions was high (Fig. 3G). Mixed-lineage cells of the same colour were occasionally detected in both nipple (Fig. 3C) and distal (Fig. 3F) ductal regions, supporting the notion that embryonic MaSCs may possess multipotent capacity (Spike et al., 2012; Van Keymeulen et al., 2011). However, a cell neighbour analysis indicated that the majority of same-colour neighbours consisted of cells of the same lineage (Fig. 3H), suggesting that embryonic stem/progenitor cells are already lineage biased in the foetal mammary gland. Indeed, a recent study, which was able to achieve multicolour labelling at clonal density, has demonstrated that Notch1-expressing cells display lineage restriction at E15.5 and E17.5 (Lilja et al., 2018). It is important to note, however, that although the results of this study at E16.5-E17.5 are consistent with previous analyses at clonal density (Lilja et al., 2018; Wuidart et al., 2018), mammary glands in our study are marked at levels higher than clonal density and thus we cannot exclude the possibility that bipotent embryonic MaSCs were initially labelled and gave rise to luminal and basal progeny that expanded only after lineage specification occurred.

A single stem cell labelled *in utero* can contribute extensively to both the basal and luminal lineages in the adult mammary gland. Whereas the low-density and neutral $R26^{CreERT2}$;$R26^{Confetti}$ model provided important corroborating evidence into the fate of primordial stem/progenitor cells to mammary development, we sought to reinforce these observations using an alternative neutral approach. To achieve this, we used the $R26^{[CA]30}$ reporter mouse model (Kozar et al., 2013) that was previously exploited to achieve unbiased, single-cell labelling in the mammary gland (Davis et al., 2016). This model encompasses a $[CA]_{30}$ microsatellite repeat positioned directly upstream of an out-of-frame modified β-glucosidase (SYNbglA) reporter gene targeted to the Rosa26 locus (Fig. 4A). During DNA replication, spontaneous frame-shift mutations in the inherently unstable dinucleotide repeat tract may place the reporter gene in-frame, leading to its expression. This ‘strand slippage’ produces a permanent mark on the cell, which is subsequently transmitted to all of its progeny. Importantly, genetic labelling in this model is exceedingly rare, thereby allowing the

Development (2018) 145, dev164079. doi:10.1242/dev.164079

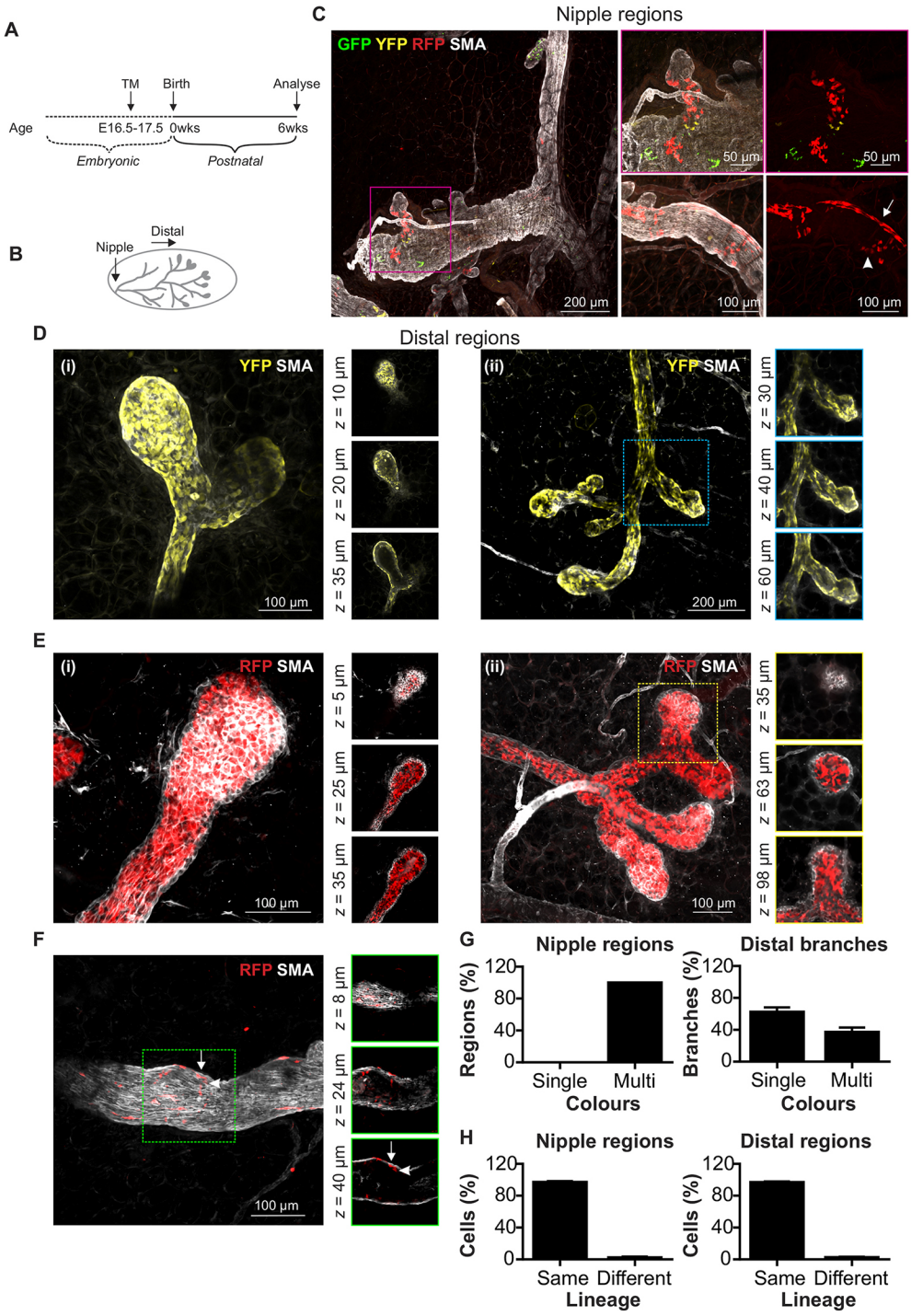


Fig. 3. Lineage tracing following embryonic labelling. (A) Tamoxifen (33 µg/g maternal weight) was administered by oral gavage (E16.5-E17.5) and tissue harvested from $R26^{CreERT2}$;$R26^{Confetti}$ offspring. (B) Analysis was divided into nipple regions and distal branches. (C) Example of a multicoloured nipple region. Arrow and arrowhead indicate adjacent RFP$^+$ basal and luminal cells, respectively. (D) Example of single-colour distal branches and terminal end buds (TEBs) comprising YFP$^+$ basal cells. i and ii are two examples showing the same thing. (E) Example of single-colour distal branches and TEBs comprising RFP$^+$ luminal cells. (F) Example of a rare distal branch containing interspersed RFP$^+$ luminal and basal cells. Images show maximum-intensity *z*-projections and optical slices of a region of interest (boxed and enlarged in the right-hand panels). Arrow and arrowhead in F show adjacent RFP$^+$ basal and luminal cells, respectively. (G) Graphs (data are mean ±s.e.m.) showing the percentage of single- and multicoloured nipple regions and distal branches. (H) Cell neighbour analysis showing that the majority of FP$^+$ cells had a same-colour FP$^+$ neighbour of the same lineage [data are mean±s.e.m. of 940 cells (seven nipple regions, n=5 mice) and 4439 cells (85 distal branches, n=7 mice) from randomly-selected 3D images].

fate of a single-labelled cell to be traced with a high degree of confidence.

Using this model, we observed variable numbers of label-positive cells randomly intermixed with unlabelled cells in developing ducts (Fig. 4B), mirroring the stochastic labelling pattern observed in pubertal $R26^{CreERT2}$;$R26^{Confetti}$ mice (Fig. 1C,D). The majority of labelled progeny arising from a single β-glucosidase$^+$ cell expressed markers of the luminal lineage (Fig. 4C), potentially reflecting the higher proliferative capacity in this compartment compared to basal cells (Giraddi et al., 2015). Isolated regions containing limited numbers of label-positive cells were most-commonly observed (Fig. 4B). These most likely arose from a recent frame-shift mutation, or from strand slippage in replicative-restricted progenitors or differentiated cells. Occasionally, large, contiguous clonal regions spanning several ductal branches were also observed, which were considered to have arisen from a single proliferative MaSC/progenitor cell.

On one occasion, we observed ductal regions that comprised exclusively β-glucosidase$^+$ basal cells in close proximity to regions comprising only β-glucosidase$^+$ luminal cells (Fig. 4D,E). The expansive size of this clone (>10 mm), in addition to its location at the nipple region of the mammary gland, suggest that a bipotent MaSC was labelled at some point during embryogenesis, giving rise to a luminal and a basal daughter cell that later generated lineage-

Development (2018) 145, dev164079. doi:10.1242/dev.164079

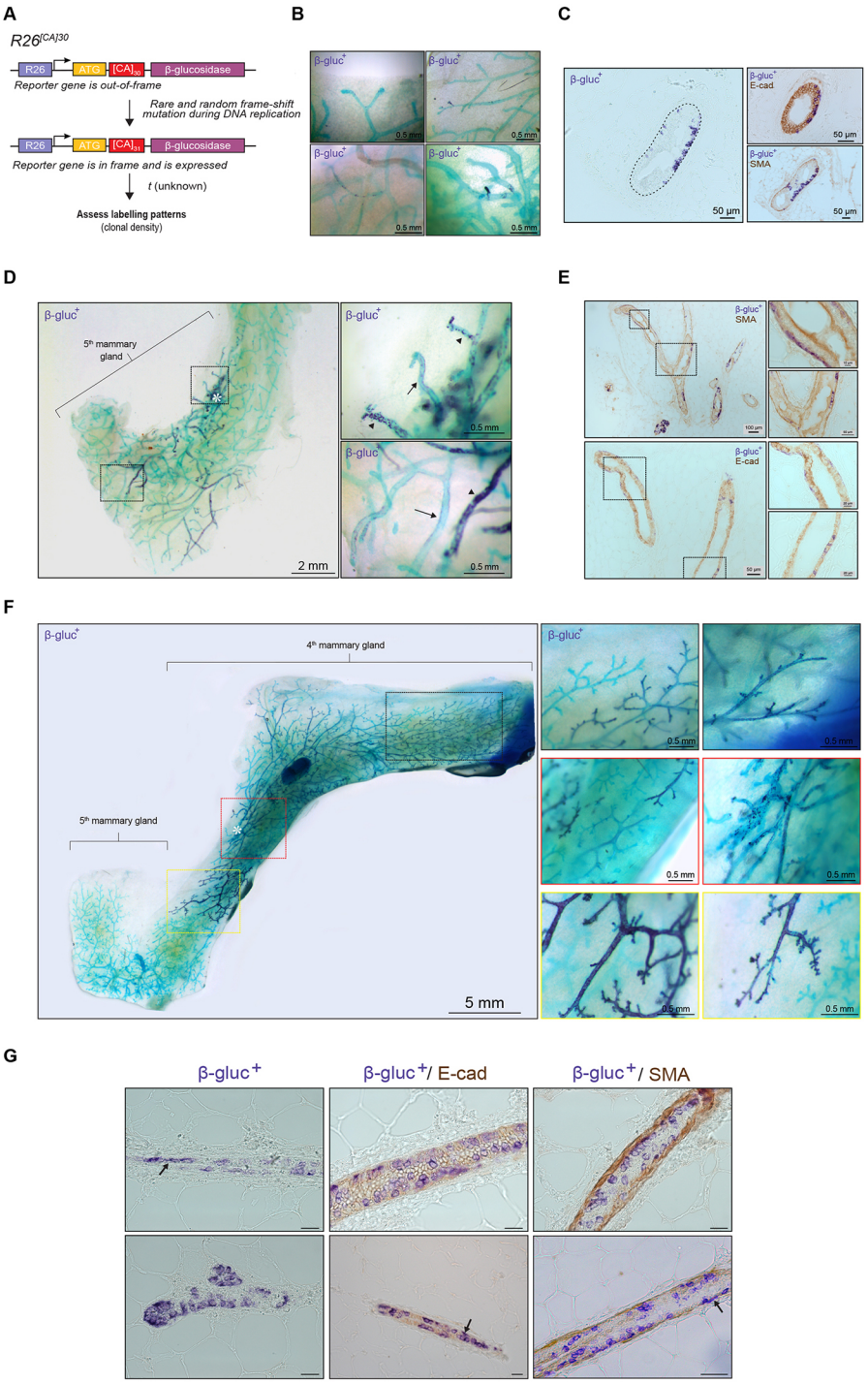


Fig. 4. Single-cell genetic labelling using the *R26*$^{[CAJ30SYNbglA}$ model. (A) Schematic representation of the *R26*$^{[CAJ30}$ model. (B,C) Dispersion of β-glucosidase^{+} cells (purple) throughout the ductal epithelium (green) (B) and lineage analysis of β-glucosidase^{+} cells with immunohistochemistry (C). (D,E) Example of a β-glucosidase^{+} clone that originated at the nipple region (asterisk) (D) and consisted of regions containing β-glucosidase^{+} luminal cells adjacent to regions containing β-glucosidase^{+} basal cells (E). Arrows and arrowheads show cells with a basal and luminal morphology, respectively. (F) A rare clone spanning the entire abdominal mammary gland. Asterisk indicates the nipple region. (G) Lineage analysis of β-glucosidase^{+} cells with immunohistochemistry. Black arrows indicate β-glucosidase^{+} basal cells among β-glucosidase^{+} luminal cells. Scale bars: 20 µm. Clones shown in D and F represent two out of three putative embryonic labelling events observed from the analysis of 30 mice.

restricted progeny during postnatal development. Remarkably, we also observed a clone that spanned the entire abdominal mammary gland (Fig. 4F). This is in stark contrast to the contribution of cells labelled during late embryogenesis in *R26*^{CreERT2}*;R26*^{Confetti} mice, which appeared, overall, more limited in scope (Fig. 3). The vast nature of this exceptional clone (and its origin at the nipple region) might, therefore, suggest that a cell was marked very early in development, likely before E12.5-E13.5, where clone size was considerably and consistently smaller (Lilja et al., 2018; van Amerongen et al., 2012; Wuidart et al., 2018). However, without knowing the precise stage of labelling in the *R26*$^{[CAJ30SYNbglA}$ model, the differential contribution of these putative embryonic cells to ductal morphogenesis in this model (Fig. 4D versus 4F) could also point to heterogeneity and multiplicity within the embryonic MaSC compartment. Histochemical analysis of this rare clone (Fig. 4F) revealed that the majority of labelled cells were luminal (Fig. 4G), further suggesting that some MaSCs might exhibit a degree of lineage bias, even during early embryonic development.

The origin of luminal and basal cell lineages in the mammary gland has been the subject of intense investigation and debate. Recent saturation lineage-tracing, single-cell lineage-tracing and promoter-driven lineage-tracing studies have provided support for lineage restriction of MaSCs from late embryogenesis into

Development (2018) 145, dev164079. doi:10.1242/dev.164079

adulthood (Davis et al., 2016; Li et al., 2016; Lilja et al., 2018; Van Keymeulen et al., 2011; Wuidart et al., 2016, 2018). However, the lack of evidence for the presence of multilineage clones does not unequivocally show that bi/multipotent stem cells do not exist (Rios et al., 2016; Visvader and Stingl, 2014). Thus, it is imperative that the epithelial hierarchy in the mammary gland is rigorously assessed at various developmental stages, using a range of methods, models and systems of analyses. Using a low-density, neutral, genetic labelling strategy and method of imaging that is free of proteolytic digestion, we have provided corroborating evidence of the lineage restriction of proliferative stem/progenitor cells to the three major stages of mammary development: in the late embryo, during puberty and in reproduction. Our findings also confirm that remarkable heterogeneity exists within the adult mammary stem and progenitor cell compartment, and suggest similar multiplicity within their embryonic precursors. Importantly, we have revealed the remarkable capacity of a single embryonic MaSC to contribute to ductal development, providing unprecedented insights that could only be disclosed by this single-cell approach. It is increasingly hypothesised that certain cancers may arise from reactivation of embryonic developmental programs in postnatal tissues (Howard and Veltmaat, 2013; Wahl and Spike, 2017). Thus, an elucidation of the full spectrum of stem/progenitor cell populations in the pre- and postnatal mammary gland is paramount for defining the cellular origin of heterogeneous breast tumours.

MATERIALS AND METHODS

Antibodies and reagents

Antibodies used in these studies include: rabbit anti-SMA (Abcam, ab5694, lot number GR248336-23, 1:200 and 1:300 for 2D and 3D studies, respectively), rat anti-K8 (Developmental Studies Hybridoma Bank, TROMA-I, 1:200 and 1:50 of supernatant for 2D and 3D studies, respectively), rabbit anti-E-cadherin (Cell Signaling, 3195, lot number 10, 1:400), goat anti-rabbit AlexaFluor (AF) 647 (Thermo Fisher Scientific, A21245, lot number 1805235, 1:500) and anti-rabbit horseradish peroxidase (HRP) (DAKO, P0448, lot number 20023997, 1:500). See supplementary Materials and Methods for further information and Lloyd-Lewis et al. (2016) for optimisation and validation studies.

Animal models

All animal experimentation was carried out in accordance with the *Animal (Scientific Procedures) Act 1986*, the *European Union Directive 86/609*, and with local ethics committee approval. Mouse (*Mus musculus*) strains $R26^{[CAJ30}$ (Kozar et al., 2013) (a kind gift from Prof. D. Winton, Cancer Research UK Cambridge Institute), $R26^{Confetti}$ (Livet et al., 2007) and $R26^{CreERT2}$ (Ventura et al., 2007) have previously been described. $R26^{[CAJ30}$ experimental mice were hemi- or homozygous for $R26^{[CAJ30SYNbglA}$. Mice were analysed for β-glucosidase expression during adulthood (7-22-weeks). Multi-colour lineage-tracing studies were performed on mice that were hemizygous for both $R26^{Confetti}$ and $R26^{CreERT2}$ ($R26^{Confetti}$;$R26^{CreERT2}$ mice). See supplementary Materials and Methods for further information.

Induction of lineage tracing in $R26^{Confetti}$;$R26^{CreERT2}$ mice

Tamoxifen was prepared in sunflower oil containing 10% ethanol. For lineage tracing during puberty in $R26^{Confetti}$;$R26^{CreERT2}$ mice, labelling was induced at the onset of puberty (4 weeks of age) by a single intraperitoneal injection of tamoxifen (0.5 mg per mouse, ∼35 µg/g) and tissue was harvested from 7-week-old mice. Using this dose, mammary gland development appeared to progress unabated, as previously reported (Rios et al., 2014). For lineage-tracing in lactating $R26^{Confetti}$;$R26^{CreERT2}$ mice, labelling was induced after puberty (12-14 weeks old) by a single intraperitoneal injection of tamoxifen (1 mg per mouse, ∼40-50 µg/g), which did not grossly affect alveolar development, as previously reported (Rios et al., 2014). After 10 days, female mice were mated with C57BL/6J male studs and lactating tissue was harvested between lactation days 4 and 5. For embryonic labelling, homozygous $R26^{Confetti}$ mice were mated with homozygous $R26^{CreERT2}$ mice. A single dose of tamoxifen (33 µg per g maternal body weight) containing progesterone (13 µg per g maternal body weight) was administered to pregnant mice via oral gavage at E16.5-17.5 (Li et al., 2016; Park et al., 2008). Using this dose, terminal end buds (TEBs) appeared morphologically normal and branching morphogenesis appeared normal via stereomicroscopy (Fig. S6). Mice were allowed to litter and tissue was collected from $R26^{Confetti}$;$R26^{CreERT2}$ offspring 6 weeks after birth.

Optical tissue clearing and whole-mount immunostaining

Fixed mammary tissue was cut into large pieces (∼15×15×2 mm) for immunostaining and tissue clearing, without any mechanical or enzymatic manipulation or microdissection. Optical tissue clearing was performed using either SeeDB (Ke et al., 2013) or a modified CUBIC (Reagent 1A) protocol (Susaki and Ueda, 2016), as previously described in detail (Lloyd-Lewis et al., 2016). Whole-mount immunostaining was performed prior to tissue clearing (SeeDB) or following immersion in CUBIC Reagent 1A, as previously described. See supplementary Materials and Methods for further information.

Confocal microscopy

Optically clear tissues were imaged in their respective refractive index matching solutions in 35 mm glass-bottom MatTek dishes. Images were acquired using a Leica TCS SP8 inverted confocal microscope with 10×/0.4 or 20×/0.75 HC PL APO objective lenses. All colours (GFP, YFP, RFP and far red) were imaged for consistency and quantification. CFP-expressing clones were under-represented and were not routinely imaged (see Fig. S2). See supplementary Materials and Methods for further information.

Whole-mount histochemistry

Detection of modified β-glucosidase expression in the mammary gland was performed as previously described (Davis et al., 2016). Briefly, excised mammary glands were fixed at room temperature for 4 h in NBF (10%). Tissue was heated to 65°C for 15 min in phosphate-buffered saline for endogenous β-glucosidase inactivation. Whole-mount mammary glands were incubated for 24 h at 50°C in a solution containing one part solution A [5-bromo-6-chloro-3-indolyl-β-D-glucopyranoside (1%) in dimethyl sulfoxide] and 25 parts solution B [magnesium chloride (0.02% w/v), potassium ferricyanide (0.096% w/v) and potassium ferrocyanide (0.13% w/v) in PBS]. After 24 h, the substrate was replenished and tissue incubated for an additional 24 h. Mammary glands were post-fixed in 10% NBF overnight at 4°C. Tissue clearing was performed using the CUBIC clearing protocol (Susaki et al., 2014), with methyl green counterstaining, as previously described (Davis et al., 2016; Lloyd-Lewis et al., 2016).

Histology

For histological analysis of tissue from $R26^{[CAJ30SYNbglA}$ mice, CUBIC-based tissue clearing was reversed by overnight incubation in PBS at 4°C. Paraffin processing was performed using a butanol clearing protocol, to maintain the histochemical magenta staining. Briefly, tissue was placed in a cassette and immersed in 70% ethanol (2 h), 96% ethanol (2 h), 100% ethanol (2 h) and finally transferred to n-butanol for 2 h before paraffin wax embedding. Paraffin wax-embedded sections (4-6 µm) were de-waxed in xylene (3×2 min washes) and processed as described above. Primary antibodies used for immunohistochemistry on paraffin slides were: rabbit anti-SMA (Abcam, ab5694, 1:200) and rabbit anti-E-cadherin (Cell Signaling, 3195, 1:50). Goat anti-rabbit HRP-conjugated secondary antibody (Jackson ImmunoResearch) was used at a dilution of 1:250.

Clonal analysis method

A cell neighbour analysis was used to analyse labelling outcomes in this study and is described in detail in the supplementary Materials and Methods. Briefly, we created *z*-projections of randomly selected 3D image stacks containing label-positive cells. For all cells within each region, the lineage of the closest same-colour neighbour was recorded as either ‘same’ or ‘different’ by manual scoring. GFP, YFP, RFP and far red channels were imaged for each image sequence.

Acknowledgements
We thank the Department of Pathology Biological Services Unit for help with animal work and H. Skelton for help with histology. All work was performed at the Cambridge Advanced Imaging Centre with support from Drs Leila Muresan, Kevin O'Holleran and Martin Lenz. We also thank Drs Ueda and Susaki (RIKEN Quantitative Biology Center) for their assistance with CUBIC-based tissue clearing.

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: B.L.-L., F.M.D., C.J.W.; Methodology: B.L.-L., F.M.D.,; Validation: B.L.-L., F.M.D., O.B.H., J.R.H.; Formal analysis: B.L.-L., F.M.D.; Investigation: B.L.-L., F.M.D., O.B.H., J.R.H.; Resources: C.J.W.; Writing - original draft: B.L.-L., F.M.D.; Writing - review & editing: B.L.-L., F.M.D., O.B.H., C.J.W.; Supervision: C.J.W.; Funding acquisition: C.J.W.

Funding
This work was supported by the Medical Research Council (MR/J001023/1 to C.J.W. and B.L.-L., and MR/N022963/1 to C.J.W. and J.R.H.], the National Health and Medical Research Council [1071074 and 1141008 to F.M.D.] and the Wellcome Trust [105377/Z/14/Z to O.B.H.]. Deposited in PMC for immediate release.

Supplementary information
Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.164079.supplemental

References

Bach, K., Pensa, S., Grzelak, M., Hadfield, J., Adams, D. J., Marioni, J. C. and Khaled, W. T. (2017). Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat. Commun.* **8**, 2128.

Bai, L. and Rohrschneider, L. R. (2010). s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes Dev.* **24**, 1882-1892.

Blaas, L., Pucci, F., Messal, H. A., Andersson, A. B., Josue Ruiz, E., Gerling, M., Douagi, I., Spencer-Dene, B., Musch, A., Mitter, R. et al. (2016). Lgr6 labels a rare population of mammary gland progenitor cells that are able to originate luminal mammary tumours. *Nat. Cell Biol.* **18**, 1346-1356.

Boras-Granic, K., Dann, P. and Wysolmerski, J. J. (2014). Embryonic cells contribute directly to the quiescent stem cell population in the adult mouse mammary gland. *Breast Cancer Res.* **16**, 487.

Cai, S., Kalisky, T., Sahoo, D., Dalerba, P., Feng, W., Lin, Y., Qian, D., Kong, A., Yu, J., Wang, F. et al. (2017). A quiescent Bcl11b high stem cell population is required for maintenance of the mammary gland. *Cell Stem Cell* **20**, 247-260.

Chang, T. H.-T., Kunasegaran, K., Tarulli, G. A., De Silva, D., Voorhoeve, P. M. and Pietersen, A. M. (2014). New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. *Breast Cancer Res.* **16**, R1.

Cowin, P. and Wysolmerski, J. (2010). Molecular mechanisms guiding embryonic mammary gland development. *Cold Spring Harb. Perspect. Biol.* **2**, a003251.

Davis, F. M., Lloyd-Lewis, B., Harris, O. B., Kozar, S., Winton, D. J., Muresan, L. and Watson, C. J. (2016). Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat. Commun.* **7**, 13053.

de Visser, K. E., Ciampricotti, M., Michalak, E. M., Tan, D. W.-M., Speksnijder, E. N., Hau, C.-S., Clevers, H., Barker, N. and Jonkers, J. (2012). Developmental stage-specific contribution of LGR5(+) cells to basal and luminal epithelial lineages in the postnatal mammary gland. *J. Pathol.* **228**, 300-309.

Fu, N. Y., Rios, A. C., Pal, B., Law, C. W., Jamieson, P., Liu, R., Vaillant, F., Jackling, F., Liu, K. H., Smyth, G. K. et al. (2017). Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive. *Nat. Cell Biol.* **19**, 164-176.

Giraddi, R. R., Shehata, M., Gallardo, M., Blasco, M. A., Simons, B. D. and Stingl, J. (2015). Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. *Nat. Commun.* **6**, 8487.

Gjorevski, N. and Nelson, C. M. (2011). Integrated morphodynamic signalling of the mammary gland. *Nat. Rev. Mol. Cell Biol.* **12**, 581-593.

Hens, J. R. and Wysolmerski, J. J. (2005). Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland. *Breast Cancer Res.* **7**, 220-224.

Hinck, L. and Silberstein, G. B. (2005). Key stages in mammary gland development: the mammary end bud as a motile organ. *Breast Cancer Res.* **7**, 245-251.

Howard, B. A. and Veltmaat, J. M. (2013). Embryonic mammary gland development; a domain of fundamental research with high relevance for breast cancer research. *J. Mammary Gland Biol. Neoplasia* **18**, 89-91.

Ke, M.-T., Fujimoto, S. and Imai, T. (2013). SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nat. Neurosci.* **16**, 1154-1161.

Kozar, S., Morrissey, E., Nicholson, A. M., van der Heijden, M., Zecchini, H. I., Kemp, R., Tavaré, S., Vermeulen, L. and Winton, D. J. (2013). Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell* **13**, 626-633.

Lafkas, D., Rodilla, V., Huyghe, M., Mourao, L., Kiaris, H. and Fre, S. (2013). Notch3 marks clonogenic mammary luminal progenitor cells in vivo. *J. Cell Biol.* **203**, 47-56.

Li, J., Miao, L., Shieh, D., Spiotto, E., Li, J., Zhou, B., Paul, A., Schwartz, R. J., Firulli, A. B., Singer, H. A. et al. (2016). Single-cell lineage tracing reveals that oriented cell division contributes to trabecular morphogenesis and regional specification. *Cell Rep.* **15**, 158-170.

Lilja, A. M., Rodilla, V., Huyghe, M., Hannezo, E., Landragin, C., Renaud, O., Leroy, A., Rulands, S., Simons, B. D. and Fre, S. (2018). Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. *Nat. Cell Biol.* **20**, 677-687.

Livet, J., Weissman, T. A., Kang, H., Draft, R. W., Lu, J., Bennis, R. A., Sanes, J. R. and Lichtman, J. W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56-62.

Lloyd-Lewis, B., Davis, F. M., Harris, O. B., Hitchcock, J. R., Lourenco, F. C., Pasche, M. and Watson, C. J. (2016). Imaging the mammary gland and mammary tumours in 3D: Optical tissue clearing and immunofluorescence methods. *Breast Cancer Res.* **18**, 127.

Lloyd-Lewis, B., Harris, O. B., Watson, C. J. and Davis, F. M. (2017). Mammary stem cells: premise, properties and perspectives. *Trends Cell Biol.* **27**, 556-567.

Paine, I., Chauviere, A., Landua, J., Sreekumar, A., Cristini, V., Rosen, J. and Lewis, M. T. (2016). A geometrically-constrained mathematical model of mammary gland ductal elongation reveals novel cellular dynamics within the terminal end bud. *PLoS Comput. Biol.* **12**, e1004839.

Park, E. J., Sun, X., Nichol, P., Saijoh, Y., Martin, J. F. and Moon, A. M. (2008). System for tamoxifen-inducible expression of Cre-recombinase from the Foxa2 locus in mice. *Dev. Dyn.* **237**, 447-453.

Prater, M. D., Petit, V., Alasdair Russell, I., Giraddi, R. R., Shehata, M., Menon, S., Schulte, R., Kalajzic, I., Rath, N., Olson, M. F. et al. (2014). Mammary stem cells have myoepithelial cell properties. *Nat. Cell Biol.* **16**, 942-950, 1–7.

Rios, A. C., Fu, N. Y., Lindeman, G. J. and Visvader, J. E. (2014). In situ identification of bipotent stem cells in the mammary gland. *Nature* **506**, 322-327.

Rios, A. C., Fu, N. Y., Cursons, J., Lindeman, G. J. and Visvader, J. E. (2016). The complexities and caveats of lineage tracing in the mammary gland. *Breast Cancer Res.* **18**, 116.

Rodilla, V., Dasti, A., Huyghe, M., Lafkas, D., Laurent, C., Reyal, F. and Fre, S. (2015). Luminal progenitors restrict their lineage potential during mammary gland development. *PLoS Biol.* **13**, e1002069.

Sale, S. and Pavelic, K. (2015). Mammary lineage tracing: the coming of age. *Cell. Mol. Life Sci.* **72**, 1577-1583.

Šale, S., Lafkas, D. and Artavanis-Tsakonas, S. (2013). Notch2 genetic fate mapping reveals two previously unrecognized mammary epithelial lineages. *Nat. Cell Biol.* **15**, 451-460.

Scheele, C. L., Hannezo, E., Muraro, M. J., Zomer, A., Langedijk, N. S., van Oudenaarden, A., Simons, B. D. and van Rheenen, J. (2017). Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* **542**, 313-317. doi:

Spike, B. T., Engle, D. D., Lin, J. C., Cheung, S. K., La, J. and Wahl, G. M. (2012). A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell Stem Cell* **10**, 183-197.

Sreekumar, A., Roarty, K. and Rosen, J. M. (2015). The mammary stem cell hierarchy: a looking glass into heterogeneous breast cancer landscapes. *Endocr. Relat. Cancer* **22**, T161-T176.

Sreekumar, A., Toneff, M. J., Toh, E., Roarty, K., Creighton, C. J., Belka, G. K., Lee, D.-K., Xu, J., Chodosh, L. A., Richards, J. S. et al. (2017). WNT-mediated regulation of FOXO1 constitutes a critical axis maintaining pubertal mammary stem cell homeostasis. *Dev. Cell* **43**, 436-448.e6.

Sun, P., Yuan, Y., Li, A., Li, B. and Dai, X. (2010). Cytokeratin expression during mouse embryonic and early postnatal mammary gland development. *Histochem. Cell Biol.* **133**, 213-221.

Susaki, E. A. and Ueda, H. R. (2016). Whole-body and whole-organ clearing and imaging techniques with single-cell resolution: toward organism-level systems biology in mammals. *Cell Chem. Biol.* **23**, 137-157.

Susaki, E. A., Tainaka, K., Perrin, D., Kishino, F., Tawara, T., Watanabe, T. M., Yokoyama, C., Onoe, H., Eguchi, M., Yamaguchi, S. et al. (2014). Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* **157**, 726-739.

Tao, L., Van Bragt, M. P. A., Laudadio, E. and Li, Z. (2014). Lineage tracing of mammary epithelial cells using cell-type-specific cre-expressing adenoviruses. *Stem Cell Rep.* **2**, 770-779.

Trejo, C. L., Luna, G., Dravis, C., Spike, B. T. and Wahl, G. M. (2017). Lgr5 is a marker for fetal mammary stem cells, but is not essential for stem cell activity or tumorigenesis. *NPJ Breast Cancer* **3**, 16.
van Amerongen, R., Bowman, A. N. N., Nusse, R., van Amerongen, R., Bowman, A. N. N. and Nusse, R. (2012). Developmental stage and time dictate the fate of Wnt/β-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* **11**, 387-400.
Van Keymeulen, A., Rocha, A. S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S. and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189-193.
Van Keymeulen, A., Fioramonti, M., Centonze, A., Bouvencourt, G., Achouri, Y. and Blanpain, C. (2017). Lineage-restricted mammary stem cells sustain the development, homeostasis, and regeneration of the estrogen receptor positive lineage. *Cell Rep.* **20**, 1525-1532.
Veltmaat, J. M. (2017). Prenatal mammary gland development in the mouse: research models and techniques for its study from past to present. *Methods Mol. Biol.* **1501**, 21-76.
Veltmaat, J. M., Mailleux, A. A., Thiery, J. P. and Bellusci, S. (2003). Mouse embryonic mammogenesis as a model for the molecular regulation of pattern formation. *Differentiation* **71**, 1-17.
Ventura, A., Kirsch, D. G., McLaughlin, M. E., Tuveson, D. A., Grimm, J., Lintault, L., Newman, J., Reczek, E. E., Weissleder, R. and Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. *Nature* **445**, 661-665.
Visvader, J. E. and Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.* **28**, 1143-1158.
Wahl, G. M. and Spike, B. T. (2017). Cell state plasticity, stem cells, EMT, and the generation of intra-tumoral heterogeneity. *NPJ Breast Cancer* **3**, 14.
Wang, D., Cai, C., Dong, X., Yu, Q. C., Zhang, X.-O., Yang, L. and Zeng, Y. A. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517**, 81-84.
Wang, C., Christin, J. R., Oktay, M. H. and Guo, W. (2017). Lineage-biased stem cells maintain estrogen-receptor-positive and -negative mouse mammary luminal lineages. *Cell Rep.* **18**, 2825-2835.
Watson, C. J. and Khaled, W. T. (2008). Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development* **135**, 995-1003.
Wuidart, A., Ousset, M., Rulands, S., Simons, B., Van Keymeulen, A. and Blanpain, C. (2016). Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes Dev.* **30**, 1261-1277.
Wuidart, A., Sifrim, A., Fioramonti, M., Matsumura, S., Brisebarre, A., Brown, D., Centonze, A., Dannau, A., Dubois, C., Van Keymeulen, A. et al. (2018). Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nat. Cell Biol.* **20**, 666-676.

References

Adler, N.T., Resko, J.A., and Goy, R.W. (1970). The effect of copulatory behavior on hormonal change in the female rat prior to implantation. Physiol. Behav. *5*: 1003–7.

Aihara, E., Mahe, M.M., Schumacher, M.A., Matthis, A.L., Feng, R., Ren, W., et al. (2015). Characterization of stem/progenitor cell cycle using murine circumvallate papilla taste bud organoid. Sci. Rep. *5*: 17185.

Aloia, L., McKie, M.A., and Huch, M. (2016). Cellular plasticity in the adult liver and stomach. J. Physiol. *594*: 4815–4825.

Amerongen, R. van, Bowman, A.N.N., Nusse, R., van Amerongen, R., Bowman, A.N.N., and Nusse, R. (2012). Developmental Stage and Time Dictate the Fate of Wnt/β-Catenin-Responsive Stem Cells in the Mammary Gland. Cell Stem Cell *11*: 387–400.

Arnaout, M.A., Todd, R.F., Dana, N., Melamed, J., Schlossman, S.F., Colten, H.R., et al. (1983). Inhibition of phagocytosis of complement C3- or immunoglobulin G-coated particles and of C3bi binding by monoclonal antibodies to a monocyte-granulocyte membrane glycoprotein (Mol). J. Clin. Invest. *72*: 171–9.

Asselin-Labat, M.-L., Shackleton, M., Stingl, J., Vaillant, F., Forrest, N.C., Eaves, C.J., et al. (2006). Steroid Hormone Receptor Status of Mouse Mammary Stem Cells. JNCI J. Natl. Cancer Inst. *98*: 1011–1014.

Asselin-Labat, M.-L., Vaillant, F., Sheridan, J.M., Pal, B., Wu, D., Simpson, E.R., et al. (2010). Control of mammary stem cell function by steroid hormone signalling. Nature *465*: 798–802.

Aupperlee, M.D., Zhao, Y., Tan, Y.S., Leipprandt, J.R., Bennett, J., Haslam, S.Z., et al. (2014). Epidermal Growth Factor Receptor (EGFR) Signaling Is a Key Mediator of Hormone-Induced Leukocyte Infiltration in the Pubertal Female Mammary Gland. Endocrinology *155*: 2301–2313.

Azaripour, A., Lagerweij, T., Scharfbillig, C., Jadczak, A.E., Willershausen, B., and Noorden, C.J.F. Van (2016). A survey of clearing techniques for 3D imaging of tissues with special reference to connective tissue. Prog. Histochem. Cytochem. *51*: 9–23.

Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell *174*: 1293–1308.e36.

Bach, K., Pensa, S., Grzelak, M., Hadfield, J., Adams, D.J., Marioni, J.C., et al. (2017). Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. Nat. Commun. *8*: 2128.

Bai, L., and Rohrschneider, L.R. (2010). s-SHIP promoter expression marks activated stem cells

in developing mouse mammary tissue. Genes Dev. *24*: 1882–92.

Bajar, B.T., Lam, A.J., Badiee, R.K., Oh, Y.-H., Chu, J., Zhou, X.X., et al. (2016). Fluorescent indicators for simultaneous reporting of all four cell cycle phases. Nat. Methods *13*: 993–996.

Barker, N., Huch, M., Kujala, P., Wetering, M. van de, Snippert, H.J., Es, J.H. van, et al. (2010). Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell *6*: 25–36.

Blanpain, C. (2013). Tracing the cellular origin of cancer. Nat. Cell Biol. *15*: 126–34.

Bolin, F.P., Preuss, L.E., Taylor, R.C., and Ference, R.J. (1989). Refractive index of some mammalian tissues using a fiber optic cladding method. Appl. Opt. *28*: 2297.

Booth, B.W., Boulanger, C.A., Anderson, L.H., and Smith, G.H. (2011). The normal mammary microenvironment suppresses the tumorigenic phenotype of mouse mammary tumor virus-neu-transformed mammary tumor cells. Oncogene *30*: 679–689.

Booth, B.W., Mack, D.L., Androutsellis-Theotokis, A., McKay, R.D.G., Boulanger, C.A., and Smith, G.H. (2008). The mammary microenvironment alters the differentiation repertoire of neural stem cells. Proc. Natl. Acad. Sci. U. S. A. *105*: 14891–6.

Boras-Granic, K., Dann, P., and Wysolmerski, J.J. (2014). Embryonic cells contribute directly to the quiescent stem cell population in the adult mouse mammary gland. Breast Cancer Res. *16*: 487.

Boulanger, C.A., Bruno, R.D., Rosu-Myles, M., and Smith, G.H. (2012). The Mouse Mammary Microenvironment Redirects Mesoderm-Derived Bone Marrow Cells to a Mammary Epithelial Progenitor Cell Fate. Stem Cells Dev. *21*: 948–954.

Boulanger, C.A., Mack, D.L., Booth, B.W., and Smith, G.H. (2007). Interaction with the mammary microenvironment redirects spermatogenic cell fate in vivo. Proc. Natl. Acad. Sci. U. S. A. *104*: 3871–6.

Boulanger, J., Kervrann, C., Bouthemy, P., Elbau, P., Sibarita, J.-B.J.-B., and Salamero, J. (2010). Patch-based nonlocal functional for denoising fluorescence microscopy image sequences. IEEE Trans. Med. Imaging *29*: 442–54.

Brady, N.J., Farrar, M.A., and Schwertfeger, K.L. (2017). STAT5 deletion in macrophages alters ductal elongation and branching during mammary gland development. Dev. Biol. *428*: 232–244.

Brisken, C., and Ataca, D. (2015). Endocrine hormones and local signals during the development of the mouse mammary gland. Wiley Interdiscip. Rev. Dev. Biol. *4*: 181–195.

Brisken, C., and O'Malley, B. (2010). Hormone Action in the Mammary Gland. Cold Spring Harb. Perspect. Biol. *2*: a003178–a003178.

Bruna, A., Rueda, O.M., Greenwood, W., Batra, A.S., Callari, M., Batra, R.N., et al. (2016). A

Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. Cell *167*: 260–274.e22.

Buczacki, S.J.A., Zecchini, H.I., Nicholson, A.M., Russell, R., Vermeulen, L., Kemp, R., et al. (2013). Intestinal label-retaining cells are secretory precursors expressing Lgr5. Nature *495*: 65–69.

Cai, C., Yu, Q.C., Jiang, W., Liu, W., Song, W., Yu, H., et al. (2014). R-spondin1 is a novel hormone mediator for mammary stem cell self-renewal. Genes Dev. *28*: 2205–2218.

Cai, S., Kalisky, T., Sahoo, D., Dalerba, P., Feng, W., Lin, Y., et al. (2017). A Quiescent Bcl11b High Stem Cell Population Is Required for Maintenance of the Mammary Gland. Cell Stem Cell *20*: 247–260.e5.

Cairns, J. (1975). Mutation selection and the natural history of cancer. Nature *255*: 197–200.

Campbell, J.J., Husmann, A., Hume, R.D., Watson, C.J., and Cameron, R.E. (2017). Development of three-dimensional collagen scaffolds with controlled architecture for cell migration studies using breast cancer cell lines. Biomaterials *114*: 34–43.

Carroll, T.D., Newton, I.P., Chen, Y., Blow, J.J., and Näthke, I. (2018). Lgr5+ intestinal stem cells reside in an unlicensed G $_1$ phase. J. Cell Biol. *217*: 1667–1685.

Cassidy, J.W. (2014). Nanotechnology in the Regeneration of Complex Tissues. Bone Tissue Regen. Insights *5*: BTRI.S12331.

Cassidy, J.W., Batra, A.S., Greenwood, W., and Bruna, A. (2016). Patient-derived tumour xenografts for breast cancer drug discovery. Endocr. Relat. Cancer *23*: T259–T270.

Cassidy, J.W., and Bruna, A. (2017). Tumor Heterogeneity. Patient Deriv. Tumor Xenograft Model. 37–55.

Chakrabarti, R., Celià-Terrassa, T., Kumar, S., Hang, X., Wei, Y., Choudhury, A., et al. (2018). Notch ligand Dll1 mediates cross-talk between mammary stem cells and the macrophageal niche. Science (80-.).

Chang, T.H.-T., Kunasegaran, K., Tarulli, G.A., Silva, D. De, Voorhoeve, P.M., and Pietersen, A.M. (2014). New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. Breast Cancer Res. *16*: R1.

Chen, F., Tillberg, P.W., and Boyden, E.S. (2015). Optical imaging. Expansion microscopy. Science *347*: 543–8.

Chen, F., Wassie, A.T., Cote, A.J., Sinha, A., Alon, S., Asano, S., et al. (2016). Nanoscale imaging of RNA with expansion microscopy. Nat. Methods *13*: 679–684.

Chen, X., Nadiarynkh, O., Plotnikov, S., and Campagnola, P.J. (2012). Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. Nat. Protoc. *7*: 654–669.

Chepko, G., and Dickson, R.B. (2003). Ultrastructure of the putative stem cell niche in rat

mammary epithelium. Tissue Cell *35*: 83–93.

Chepko, G., and Smith, G.H. (1997). Three division-competent, structurally-distinct cell populations contribute to murine mammary epithelial renewal. Tissue Cell *29*: 239–253.

Chu, E.Y., Hens, J., Andl, T., Kairo, A., Yamaguchi, T.P., Brisken, C., et al. (2004). Canonical WNT signaling promotes mammary placode development and is essential for initiation of mammary gland morphogenesis. Development *131*: 4819–29.

Chua, A.C.L., Hodson, L.J., Moldenhauer, L.M., Robertson, S.A., Ingman, W. V, Fleisch, H., et al. (2010). Dual roles for macrophages in ovarian cycle-associated development and remodelling of the mammary gland epithelium. Development *137*: 4229–38.

Chung, K., Wallace, J., Kim, S.-Y., Kalyanasundaram, S., Andalman, A.S., Davidson, T.J., et al. (2013). Structural and molecular interrogation of intact biological systems. Nature *497*: 332–7.

Clevers, H., and Watt, F.M. (2018). Defining Adult Stem Cells by Function, not by Phenotype. *13*:.

Cole, H.A. (1933). The Mammary Gland of the Mouse, during the Oestrous Cycle, Pregnancy and Lactation. Proc. R. Soc. London. Ser. B, Contain. Pap. a Biol. Character *114*: 136–161.

Cotsarelis, G., Sun, T.T., and Lavker, R.M. (1990). Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. Cell *61*: 1329–37.

Couldrey, C., Moitra, J., Vinson, C., Anver, M., Nagashima, K., and Green, J. (2002). Adipose tissue: A vital in vivo role in mammary gland development but not differentiation. Dev. Dyn. *223*: 459–468.

Cowin, P., and Wysolmerski, J. (2010). Molecular mechanisms guiding embryonic mammary gland development. Cold Spring Harb. Perspect. Biol. *2*: a003251.

Cunha, G., Young, P., Christov, K., Guzman, R., Nandi, S., Talamantes, F., et al. (1995). Mammary Phenotypic Expression Induced in Epidermal Cells by Embryonic Mammary Mesenchyme. Cells Tissues Organs *152*: 195–204.

Daniel, C.W., Ome, K.B. De, Young, J.T., Blair, P.B., and Faulkin, L.J. (1968). The in vivo life span of normal and preneoplastic mouse mammary glands: a serial transplantation study. Proc. Natl. Acad. Sci. *61*: 53–60.

Daniel, C.W., and Smith, G.H. (1999). The mammary gland: a model for development. J. Mammary Gland Biol. Neoplasia *4*: 3–8.

Davis, F.M., Lloyd-Lewis, B., Harris, O.B., Kozar, S., Winton, D.J., Muresan, L., et al. (2016). Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. Nat. Commun. *7*: 13053.

DeOme, K.B., Faulkin, L.J., Bern, H.A., and Blair, P.B. (1959). Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer Res. *19*: 515–20.

Dills, W.L. (1993). Protein fructosylation: fructose and the Maillard reaction. Am. J. Clin. Nutr. *58*: 779S–787S.

Dunbar, M.E., Dann, P.R., Robinson, G.W., Hennighausen, L., Zhang, J.P., and Wysolmerski, J.J. (1999). Parathyroid hormone-related protein signaling is necessary for sexual dimorphism during embryonic mammary development. Development *126*: 3485–93.

Elo, T., Lindfors, P.H., Lan, Q., Voutilainen, M., Trela, E., Ohlsson, C., et al. (2017). Ectodysplasin target gene Fgf20 regulates mammary bud growth and ductal invasion and branching during puberty. Sci. Rep. *7*: 5049.

Emerman, J.T., and Pitelka, D.R. (1977). Maintenance and induction of morphological differentiation in dissociated mammary epithelium on floating collagen membranes. In Vitro *13*: 316–28.

Ertürk, A., Becker, K., Jährling, N., Mauch, C.P., Hojer, C.D., Egen, J.G., et al. (2012). Three-dimensional imaging of solvent-cleared organs using 3DISCO. Nat. Protoc. *7*: 1983–1995.

Fata, J.E., Chaudhary, V., and Khokha, R. (2001). Cellular Turnover in the Mammary Gland Is Correlated with Systemic Levels of Progesterone and Not 17β-Estradiol During the Estrous Cycle. Biol. Reprod. *65*: 680–688.

Fata, J.E., Mori, H., Ewald, A.J., Zhang, H., Yao, E., Werb, Z., et al. (2007). The MAPK(ERK-1,2) pathway integrates distinct and antagonistic signals from TGFalpha and FGF7 in morphogenesis of mouse mammary epithelium. Dev. Biol. *306*: 193–207.

Fata, J.E., Werb, Z., and Bissell, M.J. (2003). Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes. Breast Cancer Res. *6*: 1.

Fatehullah, A., Tan, S.H., and Barker, N. (2016). Organoids as an in vitro model of human development and disease. Nat. Cell Biol. *18*: 246–254.

Feng, Y., Manka, D., Wagner, K.-U., and Khan, S.A. (2007). Estrogen receptor-alpha expression in the mammary epithelium is required for ductal and alveolar morphogenesis in mice. Proc. Natl. Acad. Sci. U. S. A. *104*: 14718–23.

Fernandez-Gonzalez, R., Illa-Bochaca, I., Welm, B.E., Fleisch, M.C., Werb, Z., Ortiz-de-Solorzano, C., et al. (2009). Mapping mammary gland architecture using multi-scale in situ analysis. Integr. Biol. *1*: 80–89.

Fu, N.Y., Rios, A.C., Pal, B., Law, C.W., Jamieson, P., Liu, R., et al. (2017). Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive. Nat. Cell

Biol. *19*: 164–176.

Gerdes, J., Lemke, H., Baisch, H., Wacker, H.H., Schwab, U., and Stein, H. (1984). Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. J. Immunol. *133*: 1710–5.

Giraddi, R.R., Shehata, M., Gallardo, M., Blasco, M.A., Simons, B.D., Stingl, J., et al. (2015). Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. Nat. Commun. *6*: 8487.

Gjorevski, N., and Nelson, C.M. (2011). Integrated morphodynamic signalling of the mammary gland. Nat. Rev. Mol. Cell Biol. *12*: 581–93.

Gouon-Evans, V., Lin, E.Y., and Pollard, J.W. (2002). Requirement of macrophages and eosinophils and their cytokines/chemokines for mammary gland development. Breast Cancer Res. *4*: 155–64.

Gouon-Evans, V., Rothenberg, M.E., and Pollard, J.W. (2000). Postnatal mammary gland development requires macrophages and eosinophils. Development *127*: 2269–82.

Grimm, S.L., Bu, W., Longley, M.A., Roop, D.R., Li, Y., and Rosen, J.M. (2006). Keratin 6 is not essential for mammary gland development. Breast Cancer Res. *8*: R29.

Gyorki, D.E., Asselin-Labat, M.-L., Rooijen, N. van, Lindeman, G.J., and Visvader, J.E. (2009). Resident macrophages influence stem cell activity in the mammary gland. Breast Cancer Res. *11*: R62.

Haegebarth, A., and Clevers, H. (2009). Wnt signaling, lgr5, and stem cells in the intestine and skin. Am. J. Pathol. *174*: 715–21.

Hanahan, D., Weinberg, R. a. R.A.R.A., Adams, J.M., Cory, S., Aguirre-Ghiso, J.A., Ahmed, Z., et al. (2011). Hallmarks of cancer: The next generation. Cell *144*: 646–674.

Hens, J.R., and Wysolmerski, J.J. (2005). Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland. Breast Cancer Res. *7*: 220–4.

Heuberger, B., Fitzka, I., Wasner, G., and Kratochwil, K. (1982). Induction of androgen receptor formation by epithelium-mesenchyme interaction in embryonic mouse mammary gland. Proc. Natl. Acad. Sci. U. S. A. *79*: 2957–61.

Hinck, L., and Silberstein, G.B. (2005). Key stages in mammary gland development: the mammary end bud as a motile organ. Breast Cancer Res. *7*: 245–51.

Hodson, L.J., Chua, A.C.L., Evdokiou, A., Robertson, S.A., and Ingman, W. V. (2013). Macrophage Phenotype in the Mammary Gland Fluctuates over the Course of the Estrous Cycle and Is Regulated by Ovarian Steroid Hormones. Biol. Reprod. *89*: 65–65.

Horiuchi, T., and Weller, P.F. (1997). Expression of Vascular Endothelial Growth Factor by

Human Eosinophils: Upregulation by Granulocyte Macrophage Colony-stimulating Factor and Interleukin-5. Am. J. Respir. Cell Mol. Biol. *17*: 70–77.

Hoshino, K. (1962). Morphogenesis and Growth Potentiality of Mammary Glands in Mice. I. Transplantability and Growth Potentiality of Mammary Tissue of Virgin Mice. J Natl Cancer Inst *29*: 835–851.

Howard, B.A. (2012). In the beginning: the establishment of the mammary lineage during embryogenesis. Semin. Cell Dev. Biol. *23*: 574–82.

Howard, B.A., and Gusterson, B.A. (2000). Human breast development. J. Mammary Gland Biol. Neoplasia *5*: 119–37.

Howard, B.A., and Lu, P. (2014). Stromal regulation of embryonic and postnatal mammary epithelial development and differentiation. Semin. Cell Dev. Biol. *25*: 43–51.

Howard, B.A., and Veltmaat, J.M. (2013). Embryonic mammary gland development; a domain of fundamental research with high relevance for breast cancer research. Preface. J. Mammary Gland Biol. Neoplasia *18*: 89–91.

Høyer, P.E., Lyon, H., Jakobsen, P., and Andersen, A.P. Standardized methyl green-pyronin Y procedures using pure dyes. Histochem. J. *18*: 90–4.

Hsu, Y.-C., Li, L., and Fuchs, E. (2014). Transit-amplifying cells orchestrate stem cell activity and tissue regeneration. Cell *157*: 935–49.

Huang, Q., Li, F., Liu, X., Li, W., Shi, W., Liu, F.-F., et al. (2011). Caspase 3–mediated stimulation of tumor cell repopulation during cancer radiotherapy. Nat. Med. *17*: 860–866.

Huch, M., Bonfanti, P., Boj, S.F., Sato, T., Loomans, C.J.M., Wetering, M. van de, et al. (2013a). Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. EMBO J. *32*: 2708–21.

Huch, M., Dorrell, C., Boj, S.F., Es, J.H. van, Li, V.S.W., Wetering, M. van de, et al. (2013b). In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature *494*: 247–50.

Huch, M., Gehart, H., van Boxtel, R., Hamer, K., Blokzijl, F., Verstegen, M.M.A., et al. (2014). Long-Term Culture of Genome-Stable Bipotent Stem Cells from Adult Human Liver. Cell.

Hughes, C.S., Postovit, L.M., and Lajoie, G.A. (2010). Matrigel: A complex protein mixture required for optimal growth of cell culture. Proteomics *10*: 1886–1890.

Hughes, K., and Watson, C.J. (2018). Sinus-like dilatations of the mammary milk ducts, Ki67 expression, and CD3-positive T lymphocyte infiltration, in the mammary gland of wild European rabbits during pregnancy and lactation. J. Anat. *233*: 266–273.

Hume, R.D., Berry, L., Reichelt, S., D'Angelo, M., Gomm, J., Cameron, R.E., et al. (2018a). An

Engineered Human Adipose/Collagen Model for *In Vitro* Breast Cancer Cell Migration Studies. Tissue Eng. Part A ten.tea.2017.0509.

Hume, R.D., Pensa, S., Brown, E.J., Kreuzaler, P.A., Hitchcock, J., Husmann, A., et al. (2018b). Tumour cell invasiveness and response to chemotherapeutics in adipocyte invested 3D engineered anisotropic collagen scaffolds. Sci. Rep. *8*: 12658.

Humphreys, R.C., Krajewska, M., Krnacik, S., Jaeger, R., Weiher, H., Krajewski, S., et al. (1996). Apoptosis in the terminal endbud of the murine mammary gland: a mechanism of ductal morphogenesis. Development *122*: 4013–22.

Huo, Y., and Macara, I.G. (2014). The Par3-like polarity protein Par3L is essential for mammary stem cell maintenance. Nat. Cell Biol. *16*: 526–534.

Hynes, N.E., and Watson, C.J. (2010). Mammary gland growth factors: roles in normal development and in cancer. Cold Spring Harb. Perspect. Biol. *2*: a003186.

Ingman, W. V., Wyckoff, J., Gouon-Evans, V., Condeelis, J., and Pollard, J.W. (2006). Macrophages promote collagen fibrillogenesis around terminal end buds of the developing mammary gland. Dev. Dyn. *235*: 3222–3229.

Inwald, E.C., Klinkhammer-Schalke, M., Hofstädter, F., Zeman, F., Koller, M., Gerstenhauer, M., et al. (2013). Ki-67 is a prognostic parameter in breast cancer patients: results of a large population-based cohort of a cancer registry. Breast Cancer Res. Treat. *139*: 539–52.

Ismail, P.M., Li, J., DeMayo, F.J., O'Malley, B.W., and Lydon, J.P. (2002). A Novel LacZ Reporter Mouse Reveals Complex Regulation of the Progesterone Receptor Promoter During Mammary Gland Development. Mol. Endocrinol. *16*: 2475–2489.

Ito, M., Liu, Y., Yang, Z., Nguyen, J., Liang, F., Morris, R.J., et al. (2005). Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. Nat. Med. *11*: 1351–1354.

Jamieson, P.R., Dekkers, J.F., Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2017). Derivation of a robust mouse mammary organoid system for studying tissue dynamics. Development *144*: 1065–1071.

Jardé, T., Lloyd-Lewis, B., Thomas, M., Kendrick, H., Melchor, L., Bougaret, L., et al. (2016). Wnt and Neuregulin1/ErbB signalling extends 3D culture of hormone responsive mammary organoids. Nat. Commun. *7*: 13207.

Jechlinger, M., Podsypanina, K., and Varmus, H. (2009). Regulation of transgenes in three-dimensional cultures of primary mouse mammary cells demonstrates oncogene dependence and identifies cells that survive deinduction. Genes Dev. *23*: 1677–1688.

Johnsen, S., and Widder, E.A. (1999). The physical basis of transparency in biological tissue:

Ultrastructure and the minimization of light scattering. J. Theor. Biol. *199*: 181–198.

Johnson, M.D., and Mueller, S.C. (2013). Three dimensional multiphoton imaging of fresh and whole mount developing mouse mammary glands. BMC Cancer *13*: 373.

Joshi, P.A., Jackson, H.W., Beristain, A.G., Grappa, M.A. Di, Mote, P.A., Clarke, C.L., et al. (2010). Progesterone induces adult mammary stem cell expansion. Nature *465*: 803–7.

Ke, M.-T., Fujimoto, S., and Imai, T. (2013). SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. Nat. Neurosci. *16*: 1154–61.

Kee, N., Sivalingam, S., Boonstra, R., and Wojtowicz, J.M. (2002). The utility of Ki-67 and BrdU as proliferative markers of adult neurogenesis. J. Neurosci. Methods *115*: 97–105.

Kenney, N.J., Smith, G.H., Lawrence, E., Barrett, J.C., and Salomon, D.S. (2001). Identification of Stem Cell Units in the Terminal End Bud and Duct of the Mouse Mammary Gland. J. Biomed. Biotechnol. *1*: 133–143.

Keymeulen, A. Van, Lee, M.Y., Ousset, M., Brohée, S., Rorive, S., Giraddi, R.R., et al. (2015). Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. Nature *525*: 119–123.

Keymeulen, A. Van, Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., et al. (2011). Distinct stem cells contribute to mammary gland development and maintenance. Nature *479*: 189–93.

Khokha, R., and Werb, Z. (2011). Mammary gland reprogramming: metalloproteinases couple form with function. Cold Spring Harb. Perspect. Biol. *3*: a004333.

Kogata, N., and Howard, B.A. (2013). A whole-mount immunofluorescence protocol for three-dimensional imaging of the embryonic mammary primordium. J. Mammary Gland Biol. Neoplasia *18*: 227–31.

Kordon, E.C., and Smith, G.H. (1998). An entire functional mammary gland may comprise the progeny from a single cell. Development *125*: 1921–30.

Koren, S., Reavie, L., Couto, J.P. do, Silva, D. De, Stadler, M.B., Roloff, T., et al. (2015). PIK3CAH1047R induces multipotency and multi-lineage mammary tumours. Nature *525*: 114–118.

Kozar, S., Morrissey, E., Nicholson, A.M.M., van der Heijden, M., Zecchini, H.I.I., Kemp, R., et al. (2013). Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. Cell Stem Cell *13*: 626–33.

Kratochwil, K. (1969). Organ specificity in mesenchymal induction demonstrated in the embryonic development of the mammary gland of the mouse. Dev. Biol. *20*: 46–71.

Kreuzaler, P.A., Staniszewska, A.D., Li, W., Omidvar, N., Kedjouar, B., Turkson, J., et al. (2011).

Stat3 controls lysosomal-mediated cell death in vivo. Nat. Cell Biol. *13*: 303–9.

Kubota, S.I., Takahashi, K., Nishida, J., Morishita, Y., Ehata, S., Tainaka, K., et al. (2017). Whole-Body Profiling of Cancer Metastasis with Single-Cell Resolution. Cell Rep. *20*: 236–250.

Kumar, P., Tan, Y., and Cahan, P. (2017). Understanding development and stem cells using single cell-based analyses of gene expression. Development *144*: 17–32.

Lawson, D.A., Bhakta, N.R., Kessenbrock, K., Prummel, K.D., Yu, Y., Takai, K., et al. (2015). Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature *526*: 131–135.

Le, T.T., Huff, T.B., and Cheng, J.-X. (2009). Coherent anti-Stokes Raman scattering imaging of lipids in cancer metastasis. BMC Cancer *9*: 42.

Li, F., Wu, S., Zhou, J., Sun, J., Lin, Q., Lin, H., et al. (2014). Prognostic value of Ki-67 in breast cancer patients with positive axillary lymph nodes: a retrospective cohort study. PLoS One *9*: e87264.

Li, L., and Clevers, H. (2010). Coexistence of Quiescent and Active Adult Stem Cells in Mammals. Science (80-.). *327*: 542–545.

Lilja, A.M., Rodilla, V., Huyghe, M., Hannezo, E., Landragin, C., Renaud, O., et al. (2018). Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. Nat. Cell Biol. *20*: 677–687.

Lilla, J.N., and Werb, Z. (2010). Mast cells contribute to the stromal microenvironment in mammary gland branching morphogenesis. Dev. Biol. *337*: 124–33.

Lim, E., Wu, D., Pal, B., Bouras, T., Asselin-Labat, M.-L., Vaillant, F., et al. (2010). Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res. *12*: R21.

Linnemann, J.R., Miura, H., Meixner, L.K., Irmler, M., Kloos, U.J., Hirschi, B., et al. (2015). Quantification of regenerative potential in primary human mammary epithelial cells. Development *142*: 3239–51.

Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., et al. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450*: 56–62.

Lizen, B., Claus, M., Jeannotte, L., Rijli, F.M., and Gofflot, F. (2015). Perinatal induction of Cre recombination with tamoxifen. Transgenic Res. *24*: 1065–1077.

Lloyd-Lewis, B., Davis, F.M., Harris, O.B., Hitchcock, J.R., and Watson, C.J. (2018). Neutral lineage tracing of proliferative embryonic and adult mammary stem/progenitor cells. Development *145*: dev164079.

Lloyd-Lewis, B., Harris, O.B., Watson, C.J., and Davis, F.M. (2017). Mammary Stem Cells: Premise, Properties, and Perspectives. Trends Cell Biol. *27*: 556–567.

Lopez-Garcia, C., Klein, A.M., Simons, B.D., and Winton, D.J. (2010). Intestinal stem cell replacement follows a pattern of neutral drift. Science *330*: 822–5.

Mallepell, S., Krust, A., Chambon, P., and Brisken, C. (2006). Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland. Proc. Natl. Acad. Sci. U. S. A. *103*: 2196–201.

Marx, V. (2014). Microscopy: seeing through tissue. Nat. Methods *11*: 1209–1214.

Maurer, S., Junghans, A., and Vilgis, T.A. (2012). Impact of xanthan gum, sucrose and fructose on the viscoelastic properties of agarose hydrogels. Food Hydrocoll. *29*: 298–307.

McCutcheon, S.C., Jones, K., Cumming, S.A., Kemp, R., Ireland-Zecchini, H., Saunders, J.C., et al. (2010). Characterization of a heat resistant beta-glucosidase as a new reporter in cells and mice. BMC Biol. *8*: 89.

McLean, A.C., Valenzuela, N., Fai, S., and Bennett, S.A.L. (2012). Performing vaginal lavage, crystal violet staining, and vaginal cytological evaluation for mouse estrous cycle staging identification. J. Vis. Exp. e4389.

Miller, D.H., Jin, D.X., Sokol, E.S., Cabrera, J.R., Superville, D.A., Gorelov, R.A., et al. (2018). BCL11B Drives Human Mammary Stem Cell Self-Renewal In Vitro by Inhibiting Basal Differentiation. Stem Cell Reports *10*: 1131–1145.

Mort, R.L., Ford, M.J., Sakaue-Sawano, A., Lindstrom, N.O., Casadio, A., Douglas, A.T., et al. (2014). *Fucci2a:* A bicistronic cell cycle reporter that allows Cre mediated tissue specific expression in mice. Cell Cycle *13*: 2681–2696.

Mroue, R., and Bissell, M.J. (2013). Three-dimensional cultures of mouse mammary epithelial cells. Methods Mol. Biol. *945*: 221–50.

Murakami, T.C., Mano, T., Saikawa, S., Horiguchi, S.A., Shigeta, D., Baba, K., et al. (2018). A three-dimensional single-cell-resolution whole-brain atlas using CUBIC-X expansion microscopy and tissue clearing. Nat. Neurosci. *21*: 625–637.

Navarrete, M.A.H., Maier, C.M., Falzoni, R., Quadros, L.G. de A., Lima, G.R., Baracat, E.C., et al. (2005). Assessment of the proliferative, apoptotic and cellular renovation indices of the human mammary epithelium during the follicular and luteal phases of the menstrual cycle. Breast Cancer Res. *7*: R306.

Negulescu, P.A., Krasieva, T.B., Khan, A., Kerschbaum, H.H., and Cahalan, M.D. (1996). Polarity of T Cell Shape, Motility, and Sensitivity to Antigen. Immunity *4*: 421–430.

Nguyen, A. Van, and Pollard, J.W. (2002). Colony Stimulating Factor-1 Is Required to Recruit

Macrophages into the Mammary Gland to Facilitate Mammary Ductal Outgrowth. Dev. Biol. *247*: 11–25.

Nicholson, A.M., Olpe, C., Hoyle, A., Thorsen, A.-S., Rus, T., Colombé, M., et al. (2018). Fixation and Spread of Somatic Mutations in Adult Human Colonic Epithelium. Cell Stem Cell *22*: 909–918.e8.

Nishinari, K., Takaya, T., Kohyama, K., and Watase, M. (1994). Effects of Sugars on the Gel-Sol Transition of Agarose and k -Carrageenan. In Developments in Food Engineering, (Boston, MA: Springer US), pp 108–110.

O'Brien, L.E., and Bilder, D. (2013). Beyond the Niche: Tissue-Level Coordination of Stem Cell Dynamics. Annu. Rev. Cell Dev. Biol. *29*: 107–136.

Oakes, S.R., Gallego-Ortega, D., and Ormandy, C.J. (2014). The mammary cellular hierarchy and breast cancer. Cell. Mol. Life Sci. *71*: 4301–24.

Oakes, S.R., Naylor, M.J., Asselin-Labat, M.-L., Blazek, K.D., Gardiner-Garden, M., Hilton, H.N., et al. (2008). The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev. *22*: 581–6.

Oettgen, P., Kas, K., Dube, A., Gu, X., Grall, F., Thamrongsak, U., et al. (1999). Characterization of ESE-2, a Novel ESE-1-related Ets Transcription Factor That Is Restricted to Glandular Epithelium and Differentiated Keratinocytes. J. Biol. Chem. *274*: 29439–29452.

Oftedal, O.T., and Dhouailly, D. (2013). Evo-Devo of the Mammary Gland. J. Mammary Gland Biol. Neoplasia *18*: 105–120.

Orkin, R.W., Gehron, P., McGoodwin, E.B., Martin, G.R., Valentine, T., and Swarm, R. (1977). A murine tumor producing a matrix of basement membrane. J. Exp. Med. *145*: 204–20.

Paine, I., Chauviere, A., Landua, J., Sreekumar, A., Cristini, V., Rosen, J., et al. (2016). A Geometrically-Constrained Mathematical Model of Mammary Gland Ductal Elongation Reveals Novel Cellular Dynamics within the Terminal End Bud. PLoS Comput. Biol. *12*: e1004839.

Paine, I.S., and Lewis, M.T. (2017). The Terminal End Bud: the Little Engine that Could. J. Mammary Gland Biol. Neoplasia *22*: 93–108.

Pan, C., Cai, R., Quacquarelli, F.P., Ghasemigharagoz, A., Lourbopoulos, A., Matryba, P., et al. (2016). Shrinkage-mediated imaging of entire organs and organisms using uDISCO. Nat. Methods *13*: 859–867.

Park, E.J., Sun, X., Nichol, P., Saijoh, Y., Martin, J.F., and Moon, A.M. (2008). System for tamoxifen-inducible expression of cre-recombinase from the Foxa2 locus in mice. Dev. Dyn. *237*: 447–453.

Pece, S., Tosoni, D., Confalonieri, S., Mazzarol, G., Vecchi, M., Ronzoni, S., et al. (2010). Biological and Molecular Heterogeneity of Breast Cancers Correlates with Their Cancer Stem Cell

Content. Cell *140*: 62–73.

Pfefferle, A.D., Spike, B.T., Wahl, G.M., and Perou, C.M. (2015). Luminal progenitor and fetal mammary stem cell expression features predict breast tumor response to neoadjuvant chemotherapy. Breast Cancer Res. Treat. *149*: 425–37.

Pitrone, P.G., Schindelin, J., Stuyvenberg, L., Preibisch, S., Weber, M., Eliceiri, K.W., et al. (2013). OpenSPIM: an open-access light-sheet microscopy platform. Nat. Methods *10*: 598–599.

Plaks, V., Brenot, A., Lawson, D.A., Linnemann, J.R., Kappel, E.C. Van, Wong, K.C., et al. (2013). Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell Rep. *3*: 70–8.

Potten, C.S., Hume, W.J., Reid, P., and Cairns, J. (1978). The segregation of DNA in epithelial stem cells. Cell *15*: 899–906.

Potten, C.S., and Loeffler, M. (1990). Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. Development *110*: 1001–20.

Prat, A., and Perou, C.M. (2011). Deconstructing the molecular portraits of breast cancer. Mol. Oncol. *5*: 5–23.

Prick, J., Haan, G. de, Green, A.R., and Kent, D.G. (2014). Clonal heterogeneity as a driver of disease variability in the evolution of myeloproliferative neoplasms. Exp. Hematol. *42*: 841–851.

Propper, A.Y., Howard, B.A., and Veltmaat, J.M. (2013). Prenatal Morphogenesis of Mammary Glands in Mouse and Rabbit. J. Mammary Gland Biol. Neoplasia *18*: 93–104.

Rajaram, R.D., Buric, D., Caikovski, M., Ayyanan, A., Rougemont, J., Shan, J., et al. (2015). Progesterone and Wnt4 control mammary stem cells via myoepithelial crosstalk. EMBO J. *34*: 641–652.

Reed, J.R., and Schwertfeger, K.L. (2010). Immune Cell Location and Function During Post-Natal Mammary Gland Development. J. Mammary Gland Biol. Neoplasia *15*: 329–339.

Reinherz, E.L., Hussey, R.E., and Schlossman, S.F. (1980). A monoclonal antibody blocking human T cell function. Eur. J. Immunol. *10*: 758–762.

Richards, R.G., Klotz, D.M., Walker, M.P., and DiAugustine, R.P. (2004). Mammary Gland Branching Morphogenesis Is Diminished in Mice with a Deficiency of Insulin-like Growth Factor-I (IGF-I), But Not in Mice with a Liver-Specific Deletion of IGF-I. Endocrinology *145*: 3106–3110.

Rios, A.C., Fu, N.Y., Cursons, J., Lindeman, G.J., and Visvader, J.E. (2016). The complexities and caveats of lineage tracing in the mammary gland. Breast Cancer Res. *18*: 116.

Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. Nature *506*: 322–7.

Rodilla, V., Dasti, A., Huyghe, M., Lafkas, D., Laurent, C., Reyal, F., et al. (2015). Luminal

Progenitors Restrict Their Lineage Potential during Mammary Gland Development. PLOS Biol. *13*: e1002069.

Rothenberg, M.E., and Hogan, S.P. (2006). The Eosinophil. Annu. Rev. Immunol. *24*: 147–174.

Rovero, S., Amici, A., Carlo, E. Di, Bei, R., Nanni, P., Quaglino, E., et al. (2000). DNA vaccination against rat her-2/Neu p185 more effectively inhibits carcinogenesis than transplantable carcinomas in transgenic BALB/c mice. J. Immunol. *165*: 5133–42.

Rowland, M. (1972). Influence of route of administration on drug availability. J. Pharm. Sci. *61*: 70–4.

Ruan, W., and Kleinberg, D.L. (1999). Insulin-Like Growth Factor I Is Essential for Terminal End Bud Formation and Ductal Morphogenesis during Mammary Development. Endocrinology *140*: 5075–5081.

Ruan, W., Monaco, M.E., and Kleinberg, D.L. (2005). Progesterone Stimulates Mammary Gland Ductal Morphogenesis by Synergizing with and Enhancing Insulin-Like Growth Factor-I Action. Endocrinology *146*: 1170–1178.

Sachs, N., Ligt, J. de, Kopper, O., Gogola, E., Bounova, G., Weeber, F., et al. (2018). A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell *172*: 373–386.e10.

Sakakura, T., Nishizuka, Y., and Dawe, C.J. (1976). Mesenchyme-dependent morphogenesis and epithelium-specific cytodifferentiation in mouse mammary gland. Science *194*: 1439–41.

Sakakura, T., Sakagami, Y., and Nishizuka, Y. (1982). Dual origin of mesenchymal tissues participating in mouse mammary gland embryogenesis. Dev. Biol. *91*: 202–207.

Šale, S., Lafkas, D., and Artavanis-Tsakonas, S. (2013). Notch2 genetic fate mapping reveals two previously unrecognized mammary epithelial lineages. Nat. Cell Biol. *15*: 451–60.

Sale, S., and Pavelic, K. (2015). Mammary lineage tracing: the coming of age. Cell. Mol. Life Sci. *72*: 1577–1583.

Salic, A., and Mitchison, T.J. (2008). A chemical method for fast and sensitive detection of DNA synthesis in vivo. Proc. Natl. Acad. Sci. U. S. A. *105*: 2415–20.

Santos, C.O. dos, Rebbeck, C., Rozhkova, E., Valentine, A., Samuels, A., Kadiri, L.R., et al. (2013). Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. Proc. Natl. Acad. Sci. U. S. A. *110*: 7123–30.

Sargeant, T.J., Lloyd-Lewis, B., Resemann, H.K., Ramos-Montoya, A., Skepper, J., and Watson, C.J. (2014). Stat3 controls cell death during mammary gland involution by regulating uptake of milk fat globules and lysosomal membrane permeabilization. Nat. Cell Biol. *16*: 1057–68.

Sato, T., Vries, R.G., Snippert, H.J., Wetering, M. van de, Barker, N., Stange, D.E., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature

459: 262–265.

Scheele, C.L.G.J., Hannezo, E., Muraro, M.J., Zomer, A., Langedijk, N.S.M., Oudenaarden, A. van, et al. (2017). Identity and dynamics of mammary stem cells during branching morphogenesis. Nature *542*: 313–317.

Scholzen, T., and Gerdes, J. (2000). The Ki-67 protein: from the known and the unknown. J. Cell. Physiol. *182*: 311–22.

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.-L.L., et al. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*: 84–8.

Shamir, E.R., and Ewald, A.J. (2014). Three-dimensional organotypic culture: experimental models of mammalian biology and disease. Nat. Rev. Mol. Cell Biol. *15*: 647–664.

Shehata, M., Amerongen, R. van, Zeeman, A.L., Giraddi, R.R., and Stingl, J. (2014). The influence of tamoxifen on normal mouse mammary gland homeostasis. Breast Cancer Res. *16*: 411.

Shehata, M., Teschendorff, A., Sharp, G., Novcic, N., Russell, I.A., Avril, S., et al. (2012). Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. Breast Cancer Res. *14*: R134.

Shehata, M., Waterhouse, P.D., Casey, A.E., Fang, H., Hazelwood, L., and Khokha, R. (2018). Proliferative heterogeneity of murine epithelial cells in the adult mammary gland. Commun. Biol. *1*: 111.

Simpson, G.J. (2006). The diffraction barrier broken. Nature *440*: 879–880.

Singer, S.J. (1959). Preparation of an Electron-dense Antibody Conjugate. Nature *183*: 1523–1524.

Sleeman, K.E., Kendrick, H., Ashworth, A., Isacke, C.M., and Smalley, M.J. (2006). CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. Breast Cancer Res. *8*: R7.

Sleeman, K.E., Kendrick, H., Robertson, D., Isacke, C.M., Ashworth, A., and Smalley, M.J. (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. J. Cell Biol. *176*: 19–26.

Smith, G., and Medina, D. (1988). A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. J. Cell Sci. *90*: 173–183.

Smith, G.H. (2005). Label-retaining epithelial cells in mouse mammary gland divide asymmetrically and retain their template DNA strands. Development *132*: 681–687.

Snippert, H.J., Flier, L.G. van der, Sato, T., Es, J.H. van, Born, M. van den, Kroon-Veenboer, C., et al. (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell *143*: 134–44.

Sokol, E.S., Miller, D.H., Breggia, A., Spencer, K.C., Arendt, L.M., and Gupta, P.B. (2016). Growth of human breast tissues from patient cells in 3D hydrogel scaffolds. Breast Cancer Res. *18*: 19.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., et al. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc. Natl. Acad. Sci. U. S. A. *100*: 8418–23.

Spalteholz, W. (1911). Über das Durchsichtigmachen von menschlichen und tierischen Präparaten, nebst Anhang. Über Knochenfärbung. (Leipzig: S. Hirzel).

Spike, B.T.T., Engle, D.D.D., Lin, J.C.C., Cheung, S.K.K., La, J., and Wahl, G.M.M. (2012). A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. Cell Stem Cell *10*: 183–97.

Sreekumar, A., Roarty, K., and Rosen, J.M. (2015). The mammary stem cell hierarchy: a looking glass into heterogeneous breast cancer landscapes. Endocr. Relat. Cancer *22*: T161-76.

Srinivasan, K., Strickland, P., Valdes, A., Shin, G.C., and Hinck, L. (2003). Netrin-1/Neogenin Interaction Stabilizes Multipotent Progenitor Cap Cells during Mammary Gland Morphogenesis. Dev. Cell *4*: 371–382.

Stange, D.E., Koo, B.-K., Huch, M., Sibbel, G., Basak, O., Lyubimova, A., et al. (2013). Differentiated Troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. Cell *155*: 357–68.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., et al. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*: 993–7.

Stull, M.A., Pai, V., Vomachka, A.J., Marshall, A.M., Jacob, G.A., and Horseman, N.D. (2007). Mammary gland homeostasis employs serotonergic regulation of epithelial tight junctions. Proc. Natl. Acad. Sci. U. S. A. *104*: 16708–13.

Sun, P., Yuan, Y., Li, A., Li, B., and Dai, X. (2010). Cytokeratin expression during mouse embryonic and early postnatal mammary gland development. Histochem. Cell Biol. *133*: 213–21.

Susaki, E.A., Tainaka, K., Perrin, D., Kishino, F., Tawara, T., Watanabe, T.M., et al. (2014). Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. Cell *157*: 726–39.

Susaki, E.A., and Ueda, H.R. (2016). Whole-body and Whole-Organ Clearing and Imaging Techniques with Single-Cell Resolution: Toward Organism-Level Systems Biology in Mammals. Cell Chem. Biol. *23*: 137–157.

Sznurkowska, M.K., Hannezo, E., Azzarelli, R., Rulands, S., Nestorowa, S., Hindley, C.J., et al. (2018). Defining Lineage Potential and Fate Behavior of Precursors during Pancreas Development. Dev. Cell *46*: 360–375.e5.

Tainaka, K., Murakami, T.C., Susaki, E.A., Shimizu, C., Saito, R., Takahashi, K., et al. (2018). Chemical Landscape for Tissue Clearing Based on Hydrophilic Reagents. Cell Rep. *24*: 2196–2210.e9.

Taylor, R.A., Wang, H., Wilkinson, S.E., Richards, M.G., Britt, K., Vaillant, F., et al. (2009). Lineage Enforcement by Inductive Mesenchyme on Adult Epithelial Stem Cells across Developmental Germ Layers. Stem Cells N/A-N/A.

Tiede, B., and Kang, Y. (2011). From milk to malignancy: the role of mammary stem cells in development, pregnancy and breast cancer. Cell Res. *21*: 245–57.

Tsai, P.S., Kaufhold, J.P., Blinder, P., Friedman, B., Drew, P.J., Karten, H.J., et al. (2009). Correlations of neuronal and microvascular densities in murine cortex revealed by direct counting and colocalization of nuclei and vessels. J. Neurosci. *29*: 14553–70.

Tuchin, V. V. (2015). Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis.

Tumbar, T., Guasch, G., Greco, V., Blanpain, C., Lowry, W.E., Rendl, M., et al. (2004). Defining the epithelial stem cell niche in skin. Science *303*: 359–63.

Ursin, G., Astrahan, M.A., Salane, M., Parisky, Y.R., Pearce, J.G., Daniels, J.R., et al. (1998). The detection of changes in mammographic densities. Cancer Epidemiol. Biomarkers Prev. *7*: 43–7.

van Amerongen, R., Bowman, A.N., and Nusse, R. (2012). Developmental Stage and Time Dictate the Fate of Wnt/β-Catenin-Responsive Stem Cells in the Mammary Gland. Cell Stem Cell *11*: 387–400.

van de Wetering, M., Francies, H.E., Francis, J.M., Bounova, G., Iorio, F., Pronk, A., et al. (2015). Prospective Derivation of a Living Organoid Biobank of Colorectal Cancer Patients. Cell *161*: 933–945.

Veltmaat, J.M., Mailleux, A.A., Thiery, J.P., and Bellusci, S. (2003). Mouse embryonic mammogenesis as a model for the molecular regulation of pattern formation. Differentiation *71*: 1–17.

Veltmaat, J.M., Relaix, F., Le, L.T., Kratochwil, K., Sala, F.G., Veelen, W. van, et al. (2006). Gli3-mediated somitic Fgf10 expression gradients arerequired for the induction and patterning of mammary epithelium along theembryonic axes. Development *133*: 2325–2335.

Veltmaat, J.M., Veelen, W. Van, Thiery, J.P., and Bellusci, S. (2004). Identification of the mammary line in mouse byWnt10b expression. Dev. Dyn. *229*: 349–356.

Ventura, A., Kirsch, D.G., McLaughlin, M.E., Tuveson, D.A., Grimm, J., Lintault, L., et al. (2007). Restoration of p53 function leads to tumour regression in vivo. Nature *445*: 661–5.

Visser, K.E. de, Ciampricotti, M., Michalak, E.M., Tan, D.W.-M., Speksnijder, E.N., Hau, C.-S., et

al. (2012). Developmental stage-specific contribution of LGR5(+) cells to basal and luminal epithelial lineages in the postnatal mammary gland. J. Pathol. *228*: 300–9.

Visvader, J.E. (2009). Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev. *23*: 2563–77.

Visvader, J.E., and Clevers, H. (2016). Tissue-specific designs of stem cell hierarchies. Nat. Cell Biol. *18*: 349–355.

Visvader, J.E., and Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. Genes Dev. *28*: 1143–58.

Wagner, K.-U., Boulanger, C.A., Henry, M.D., Sgagias, M., Hennighausen, L., and Smith, G.H. (2002). An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. Development *129*: 1377–86.

Wahl, G.M., and Spike, B.T. (2017). Cell state plasticity, stem cells, EMT, and the generation of intra-tumoral heterogeneity. Npj Breast Cancer *3*: 14.

Wang, D., Cai, C., Dong, X., Yu, Q.C., Zhang, X.-O., Yang, L., et al. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. Nature *517*: 81–4.

Wang, H., Naghavi, M., Allen, C., Barber, R.M., Bhutta, Z.A., Carter, A., et al. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet *388*: 1459–1544.

Wansbury, O., Mackay, A., Kogata, N., Mitsopoulos, C., Kendrick, H., Davidson, K., et al. (2011). Transcriptome analysis of embryonic mammary cells reveals insights into mammary lineage establishment. Breast Cancer Res. *13*: R79.

Watson, C.J. (2006). Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. Breast Cancer Res. *8*: 203.

Watson, C.J., and Khaled, W.T. (2008). Mammary development in the embryo and adult: a journey of morphogenesis and commitment. Development *135*: 995–1003.

Williams, J.M., and Daniel, C.W. (1983). Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. Dev. Biol. *97*: 274–90.

Wuidart, A., Ousset, M., Rulands, S., Simons, B.D., Keymeulen, A. Van, and Blanpain, C. (2016). Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. Genes Dev. *30*: 1261–77.

Wuidart, A., Sifrim, A., Fioramonti, M., Matsumura, S., Brisebarre, A., Brown, D., et al. (2018). Early lineage segregation of multipotent embryonic mammary gland progenitors. Nat. Cell Biol. *20*: 666–676.

Yamaguchi, T.P., Bradley, A., McMahon, A.P., and Jones, S. (1999). A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. Development *126*: 1211–23.

Yang, B., Treweek, J.B., Kulkarni, R.P., Deverman, B.E., Chen, C.-K., Lubeck, E., et al. (2014). Single-Cell Phenotyping within Transparent Intact Tissue through Whole-Body Clearing. Cell *158*: 945–58.

Yang, J.J., Larsen, C.M., Grattan, D.R., and Erskine, M.S. (2009). Mating-induced neuroendocrine responses during pseudopregnancy in the female mouse. J. Neuroendocrinol. *21*: 30–9.

Ye, T., Xu, B., Wang, Z., Zhang, T.-Y., Hu, C.-Y., Lin, L., et al. (2014). Comparison of iodinated trihalomethanes formation during aqueous chlor(am)ination of different iodinated X-ray contrast media compounds in the presence of natural organic matter. Water Res. *66*: 390–398.

Young, L.J.T.J., Medina, D., DeOme, K.B.B., and Daniel, C.W.W. (1971). The influence of host and tissue age on life span and growth rate of serially transplanted mouse mammary gland. Exp. Gerontol. *6*: 49–56.

Zeng, C., Pan, F., Jones, L.A., Lim, M.M., Griffin, E.A., Sheline, Y.I., et al. (2010). Evaluation of 5-ethynyl-2'-deoxyuridine staining as a sensitive and reliable method for studying cell proliferation in the adult nervous system. Brain Res. *1319*: 21–32.

Zeps, N., Bentel, J.M., Papadimitriou, J.M., and Dawkins, H.J. (1999). Murine progesterone receptor expression in proliferating mammary epithelial cells during normal pubertal development and adult estrous cycle. Association with eralpha and erbeta status. J. Histochem. Cytochem. *47*: 1323–30.

Zomer, A., Ellenbroek, S.I.J., Ritsma, L., Beerling, E., Vrisekoop, N., and Rheenen, J. Van (2013). Brief Report: Intravital Imaging of Cancer Stem Cell Plasticity in Mammary Tumors. Stem Cells *31*: 602–606.

Zysk, A.M., Chaney, E.J., and Boppart, S.A. (2006). Refractive index of carcinogen-induced rat mammary tumours. Phys. Med. Biol. *51*: 2165–2177.